NAVAL POSTGRADUATE SCHOOL

Monterey, California

THESIS

NUMBER OF TEST SAMPLES NEEDED TO OBTAIN
A DESIRED BAYESIAN CONFIDENCE INTERVAL
FOR A PROPORTION

by

Ahmet Ziyaeddin Ipekkan

March 1989

Thesis Advisor: Glenn F. Lindsay

Unclassified

security classification of this page

REPORT DOCUMENTATION PAGE

1a Report Security Classification Unclassified	1b Restrictive Markings
2a Security Classification Authority	3 Distribution/Availability of Report Approved for public release; distribution is unlimited.
2b Declassification/Downgrading Schedule	
4 Performing Organization Report Number(s)	5 Monitoring Organization Report Number(s)
6a Name of Performing Organization Naval Postgraduate School; 6b Office Symbol (if applicable) 55	7a Name of Monitoring Organization Naval Postgraduate School
6c Address (city, state, and ZIP code) Monterey, CA 93943-5000	7b Address (city, state, and ZIP code) Monterey, CA 93943-5000
8a Name of Funding/Sponsoring Organization; 8b Office Symbol (if applicable)	9 Procurement Instrument Identification Number
8c Address (city, state, and ZIP code)	10 Source of Funding Numbers: Program Element No; Project No; Task No; Work Unit Accession No

11 Title (include security classification) NUMBER OF TEST SAMPLES NEEDED TO OBTAIN A DESIRED BAYESIAN CONFIDENCE INTERVAL FOR A PROPORTION

12 Personal Author(s) Ahmet Ziyaeddin IPEKKAN

13a Type of Report	13b Time Covered	14 Date of Report (year, month, day)	15 Page Count
Master's Thesis	From To	March 1989	85

16 Supplementary Notation The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.

17 Cosati Codes

Field	Group	Subgroup

18 Subject Terms (continue on reverse if necessary and identify by block number) Proportion, Prior Distribution, Bayesian Methods, Confidence Interval Size, Triangular Distribution, Posterior Dist.

19 Abstract (continue on reverse if necessary and identify by block number)

One recurring problem in military operational test and evaluation is determination of the number of items to test. This thesis describes a Bayesian method to determine the sample size that is needed to estimate a proportion or probability with a $(1-\alpha)100$ confidence when a prior distribution is given to that proportion. It uses the two variants of the triangular distribution as priors and develops computer programs, graphs, and tables to assist in finding the required sample size. These results are compared with other approaches in determining the required sample sizes that are needed to obtain a desired confidence interval for a proportion or probability.

20 Distribution/Availability of Abstract ☒ unclassified/unlimited ☐ same as report ☐ DTIC users	21 Abstract Security Classification Unclassified
22a Name of Responsible Individual Glenn F. Lindsay	22b Telephone (include Area code) (408) 646-2688; 22c Office Symbol 55Ls

DD FORM 1473,84 MAR — 83 APR edition may be used until exhausted. All other editions are obsolete — security classification of this page

Unclassified

Approved for public release; distribution is unlimited.

NUMBER OF TEST SAMPLES NEEDED TO OBTAIN A DESIRED BAYESIAN CONFIDENCE INTERVAL FOR A PROPORTION

by

Ahmet Ziyaeddin IPEKKAN
1st Lieutenant , Turkish Army
B.S., Turkish Military Academy, 1981

Submitted in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE IN OPERATIONS RESEARCH

from the

NAVAL POSTGRADUATE SCHOOL
March 1989

Author: ______________________
Ahmet Ziyaeddin IPEKKAN

Approved by: ______________________
Glenn F. Lindsay, Thesis Advisor

____
Dan C. Boger, Second Reader

____
Peter Purdue, Chairman,
Department of Operations Research

____
Kneale T. Marshall,
Dean of Information and Policy Sciences

ABSTRACT

One recurring problem in military operational test and evaluation is determination of the number of items to test. This thesis describes a Bayesian method to determine the sample size that is needed to estimate a proportion or probability with a $(1-\alpha)100$ confidence when a prior distribution is given to that proportion. It uses the two variants of the triangular distribution as priors and develops computer programs, graphs, and tables to assist in finding the required sample size. These results are compared with other approaches in determining the required sample sizes that are needed to obtain a desired confidence interval for a proportion or probability.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

I. INTRODUCTION

In the planning of a sample survey, or in many forms of weapon system testing, a stage is always reached at which a decision must be made about the size of the sample. The decision is important. Too large a sample implies a waste of resources, and too small a sample diminishes the utility of the results. The decision cannot always be made satisfactorily, for often we do not possess enough information to be sure that our choice of sample size is the best one. The topic of this thesis is to respond the question of how many observations are necessary for a given degree of accuracy, or how large the sample size should be to estimate proportions from a set of Bernoulli trials.

There are several ways to obtain estimates for unknown parameters. In this thesis, we will use the definition of a confidence interval to estimate the unknown probability or proportion. "A confidence interval for an unknown parameter gives an indication of the numerical value of our unknown parameter as well as a measure of how confident we are of that numerical value. "[Ref. 1: p. 323] Generally, the bigger the sample size used, the shorter the confidence interval will be.

Our major focus throughout this work is to determine the number of samples that are needed to produce a desired confidence interval size for a proportion or probability. This study investigates the necessary sample size that would be used with Bayesian statistical methods that make use of the existing experience of the experimenter and his knowledge of the phenomenon being studied. The uniform density and beta density functions were used as the prior distributions in [Ref. 2] and [Ref. 3] where the sample size question based on Bayesian confidence intervals was also studied. The uniform distribution on the interval (0,1) does not provide a great deal of flexibility in choosing a prior, but, it distributes our ignorance equally. The beta distribution with various parameter values allows a better control of the decision maker's prior beliefs and the representation of skewing, but it is

difficult to translate the decision maker's knowledge and judgement into the distribution parameters. In this work we will use the two variants of the triangular density function as our prior distributions, because they allow a simple representation of distributions which are either more heavily weighted in favor of high values of proportions rather than low values or low values of proportions rather than high values. After developing the relationship between needed sample size and the decision maker's prior information represented by a triangular distribution, we will provide graphs and tables to assist a decision maker in finding the number of samples needed to produce a desired confidence interval to estimate a proportion or probability.

We will begin by discussing various methods that can be used to find the number of samples needed, and we will compare these methods. In Chapter II we will describe a method to determine the sample size using classical statistics. In the next chapter we will describe our Bayesian method with the prior, sampling, and posterior distributions in order to find the sample size to estimate a proportion. We will use the two variants of the triangular density function as our prior distributions and the binomial distribution as our sampling distribution. Also, in Chapter III we will give the derivation of the posterior distributions. Then, in Chapter IV we will discuss how we developed and how we can use computer programs, graphs, and tables to determine the required sample size to obtain a desired 95% confidence interval for a probability or proportion. Also, we will explain the computer programs used for the Bayesian results.

In the final chapter we will summarize our work, and we will give some suggestions for further research.

II. DETERMINING THE DESIRED SAMPLE SIZE FOR PROPORTIONS USING THE CLASSICAL METHOD

In this chapter, we will describe how we can use classical methods to determine the desired sample size to estimate proportions.

Statistical methods are concerned with using the numbers observed in a sample from the population to make inferences about the population or, more specifica'`` a probability measure of the population. We will study estimation methods in this work. Problems of estimation are concerned with calculations of the numbers that occur in a sample to guess or estimate the values of unknown parameters of the population probability law. First we will find a point estimate for proportions. Then we will use this point estimate to find a confidence interval which provides an indication of a precision or accuracy of an estimate. Next, we will use this confidence interval to determine the required sample size.

A. POINT ESTIMATE FOR PROPORTIONS

"A point estimate of an unknown parameter is a number, computed from observed sample values, that is used as our guess for the value of the unknown parameter." [Ref. 4: p. 175] To estimate this value, we could use any number we like, if we can in some way base the number we choose on the results observed in selecting a random sample from the population.

In our case, we could calculate a point estimate, X/n, for the binomial parameter p, where X is the number of successes in n independent Bernoulli trials, each having p as the probability of success. That is,

$$\hat{p} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{X}{n} . \qquad (2.1)$$

This is the point estimate for a proportion p. Then we will use this point estimate to establish a confidence interval for the proportion.

B. THE CONFIDENCE INTERVAL FOR PROPORTIONS

When n is large, we can use the normal approximation to the binomial. That is, for n large and $np > 5$, $\hat{p}$ has approximately a normal distribution with mean p and variance $p(1-p)/n$. Since n is large, we can approximate the variance by $\hat{p}(1-\hat{p})/n$. The problem of a point estimate is converted into that of finding the confidence intervals for the mean of the normal distribution with a known variance. Thus a (1-α)100 percent large-sample confidence interval for P is given by [Ref. 5: p. 325]

$$\left(\hat{p} - Z_{1-\frac{\alpha}{2}}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}, \hat{p} + Z_{1-\frac{\alpha}{2}}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right). \tag{2.2}$$

For example, suppose that the number of defective items in a sample of 100 is 10. Then from Equation 2.1 the point estimate is $\hat{p} = \frac{10}{100} = 0.1$. A confidence interval of 90% for a proportion is

$$\left(0.1 - 1.645 \times \sqrt{\frac{0.1 \times 0.9}{100}}, \quad 0.1 + 1.645 \times \sqrt{\frac{0.1 \times 0.9}{100}}\right) = (0.05, 0.15).$$

This says that as a result of our sample of 100 items, we are 90% certain that this confidence interval (0.05, 0.15) contains the true value of our proportion. But even when this is done, the desired accuracy cannot be guaranteed. "With 95 percent confidence" or "with 90 percent confidence" means that the confidence intervals computed will in 5 or 10 cases out of 100 not include the population parameter: in these cases, the desired accuracy will not be attained. The only way of guaranteeing the stated accuracy is to measure each item in the whole population.

C. SAMPLE-SIZE DETERMINATION FOR ESTIMATING PROPORTIONS USING CONFIDENCE INTERVALS

Statisticians are often asked the question, "How many observations should I take?" Before this question can be answered, we must know what the

problem is and what kind of risks the decision maker is willing to take. Suppose the decision maker is interested in finding a $(1-\alpha)$ level confidence interval of a proportion. Suppose further that he wants a proportion to be estimated within A units of the true proportion. In that case, the length of the confidence interval desired is 2A. Then our true proportion P will be bounded by

$$\hat{p} - A \leq P \leq \hat{p} + A,$$

where A is the maximum error of estimate and

$$A = Z_{1-\frac{\alpha}{2}} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}. \tag{2.3}$$

Hence, we can determine the sample size n by solving Equation 2.3, obtaining

$$n = Z^2_{1-\frac{\alpha}{2}} \, \hat{p} \, \frac{(1-\hat{p})}{A^2}. \tag{2.4}$$

Equation 2.4 gives the value of n, and a nearest integer would suffice as the sample size. Note that the required n depends on the value of p which we must assume before we select the sample from which we will estimate it. We cannot use the actual value of $\hat{p}$ in the calculation because it is still unknown.

The sample size n given by Equation 2.4 is maximized at $p = 0.5$. Thus for worst case planning Equation 2.4 yields, for $\alpha = 0.05$,

$$n = \left(\frac{1.96}{A}\right)^2 (0.5)\ (0.5) = \frac{0.9604}{A^2}.$$

Here, for example, if we wish the size 2A of the 95% confidence interval to be 0.10, then we can find n = 385.

Values of n given by Equation 2.4 to determine different 95% confidence interval for different proportions can be found by using Table 1. For example, suppose that $\hat{p} = 0.5$, and $2A = 0.20$ is desired for a 95% confidence interval. Then we can find $n = 97$ by Table 1.

Table 1. DESIRED SAMPLE SIZE FOR 95% CONFIDENCE INTERVAL

95% Confidence Interval Size = 2A	Experimenter's Guess about Probability of Success						
	0.5	0.6	0.666	0.7	0.8	0.9	0.975
0.05	1,537	1,476	1,366	1,291	984	554	150
0.10	385	367	341	323	246	139	38
0.15	171	164	152	144	110	62	17
0.20	97	93	86	81	62	35	10
0.25	62	60	55	52	40	23	6
0.30	43	41	38	36	28	16	5

An alternative to this approach to finding sample size employs a prior distribution for P and Bayes' theorem. In the next chapter we will describe the prior distributions, the binomial sampling distribution, and posterior distributions as they are related by Bayes' theorem. We will also explain our rationale for selecting the two variants of the triangular distribution as our prior distributions.

III. A BAYESIAN METHOD FOR THE ESTIMATION OF PROPORTIONS, USING TRIANGULAR PRIORS

Statistical inference and decision problems about proportions can be dealt with using Bayesian analysis. In the Bayesian approach to statistics, an attempt is made to utilize all available information (both sample and prior information) in order to reduce the amount of uncertainty present in an inferential or decision making problem. As new information is obtained, it is combined with any previous information to form the basis for statistical procedures. The formal mechanism used to combine the new information with the previously available information is known as Bayes' theorem.

In Chapter II we discussed the use of the classical method to estimate proportions. In this chapter we will study the three parts of a Bayesian method to determine the required sample size to estimate proportions: the prior distribution, the sampling distribution and the posterior distribution. The terms "prior" and "posterior" are relative to the observed information. Next we will discuss why we selected the triangular density function as our prior distribution and the binomial as our sampling distribution. Finally, we will derive our posterior distributions by using Bayes' theorem.

A. BAYES' THEOREM

We will use Bayes' theorem throughout this work to estimate a proportion. Bayes' theorem is a relatively minor extension of the definition of conditional probability.

The typical phrasing of Bayes' theorem is in terms of disjoint events $A_1, A_2, \ldots, A_n$, whose union has probability one (i.e., one of the A_i is certain to occur). Prior probabilities $P(A_i)$, for the events, are assume known. An event B occurs, for which $P(B \mid A_i)$ (the conditional probability of B given A_i) is known for each A_i. Bayes' theorem then states that

$$P(A_i \mid B) = \frac{P(B \mid A_i)\ P(A_i)}{\sum_{j=1}^{n} P(B \mid A_j)\ P(A_j)},$$

for any $1 \leq i \leq n$. These probabilities reflect our revised opinions about the A_i, in the light of the knowledge that B has occurred. [Ref. 6: p. 129]

The version which we will use is exactly analogous if we adopt the following changes:

1. Replace $P(A_i)$ with a probability density function, f(p).
2. Replace summation, $\sum$, with integration, $\int$, and
3. Let X be the number of successes in n independent Bernoulli trials instead of B.

The version of Bayes' theorem [Ref. 7: p. 220] that results is

$$f(p \mid X) = \frac{f(X \mid P = p)\ f(p)}{\int f(X \mid P = p)\ f(p)\ dp}, \tag{3.1}$$

where P is a continuous random variable with density function f(p) so that

$$\int_{-\infty}^{\infty} f(p)\ dp = 1.$$

We will look at Equation 3.1 in three parts. The density function $f(p)$ represents the prior distribution (before any sampling), which will be one of the variants of the triangular density function in our work. The sampling function is $f(X \mid P = p)$ that is binomial with the sample size n. The probability function $f(p \mid X)$ is called the posterior distribution, obtained by combining the prior information $f(p)$ and the sample information (X). At the same time, the mean of the posterior distribution is called the Bayesian estimate of P.

In our continuous case, Bayes's theorem can be expressed in words as

$$\textit{Posterior distribution} = \frac{(\textit{Prior distribution})\ \ (\textit{Sampling distribution})}{\int (\textit{Prior distribution})\ (\textit{Sampling distribution})}.$$

B. THE SELECTION AND USE OF PRIOR DISTRIBUTIONS

The question is: How should one use the important prior information? Bayesian analysis allows effective use of prior information through statement of a prior distribution. The prior distribution should, of course, reflect the decision maker's prior information, which may occur in a wide variety of states. Larson says:

> The prior distribution of a parameter P can be a probability function or probability density function expressing our degree of belief about the value of P, prior to observing a sample of a random variable X whose distribution function depends on P.[Ref. 1: p. 553]

Different distribution functions can be characterized as "priors". In this study we will use the two variants of the triangular density function, which are developed in the Appendix A, as our priors. They are as follows:

$$f_1(p) = \begin{cases} \dfrac{2}{p_{\max}^2} \ (p_{\max} - p), & \text{for } \ 0 < p \le p_{\max} \le 1, \\ 0, & \text{elsewhere}, \end{cases} \tag{3.2}$$

and

$$f_2(p) = \begin{cases} -\dfrac{2 \ (p_{\min} - p)}{(p_{\min} - 1)^2}, & \text{for } \ 0 \le p_{\min} \le p < 1, \\ 0, & \text{elsewhere}. \end{cases} \tag{3.3}$$

We also derived the means of these two variants of the triangular density function in the Appendix A, which are respectively as follows:

$$E_1(P) = \frac{p_{\max}}{3},$$

and

$$E_2(P) = \frac{1}{3}(p_{\min} + 2).$$

These two variants of the triangular density functions are graphed as functions of P in Figures 1 and 2 respectively. We also notice the symmetry between the functions shown in Figures 1 and 2.

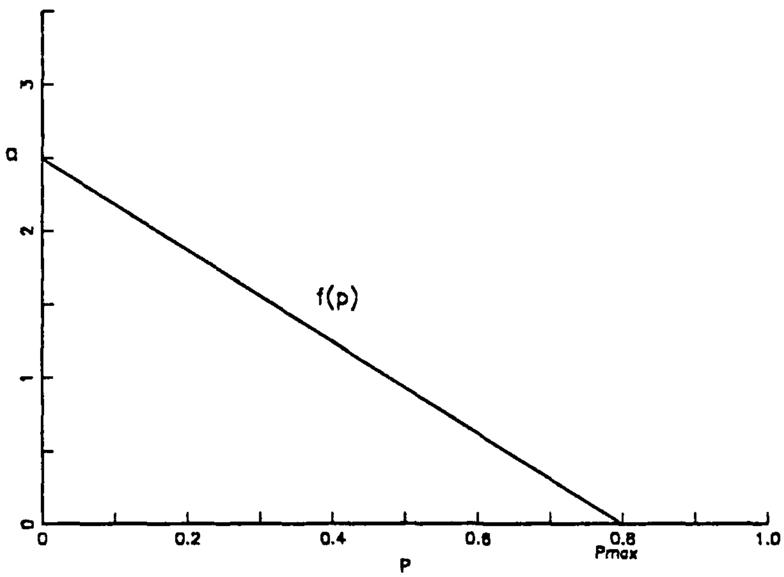

Figure 1. Triangular Density Function with Parameter (p_{max})

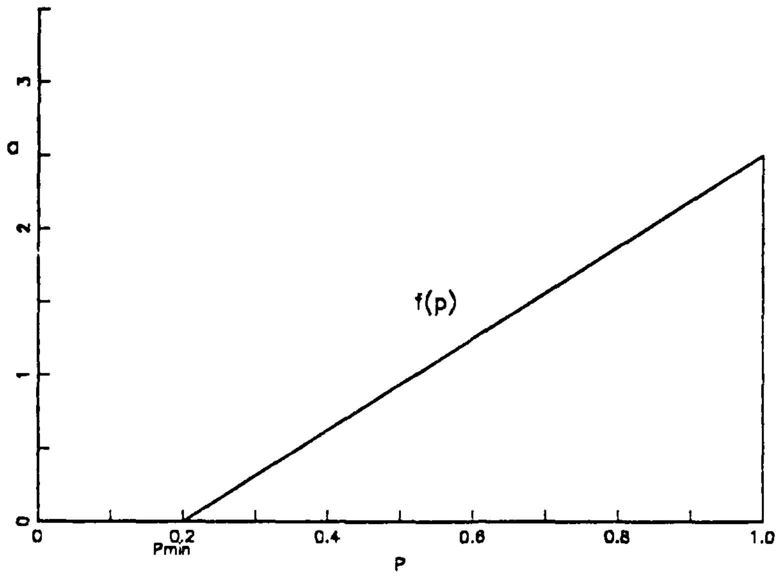

Figure 2. Triangular Density Function with Parameter (p_{min})

Note that when the triangular density functions have parameters $p_{max} = 1$ and $p_{min} = 0$ in Equation 3.2 and 3.3 respectively, they are special cases of the beta density function when $\alpha = 1$, $\beta = 2$ and when $\alpha = 2$, $\beta = 1$. They are as follows:

$$f(p \mid \alpha = 1, \beta = 2) = \begin{cases} 2\ (1 - p), & 0 \le p \le 1, \\ 0, & \mathit{elsewhere}. \end{cases} \tag{3.4}$$

and

$$f(p \mid \alpha = 2, \beta = 1) = \begin{cases} 2\,p, & 0 \le p \le 1, \\ 0, & elsewhere. \end{cases} \tag{3.5}$$

In our study the sample information will be represented in terms of a sampling function by the binomial distribution. A random variable X has a binomial distribution with parameters n and p if X has a discrete distribution for which the probability function is as follows:

$$f(x \mid n, p) = \begin{cases} \binom{n}{x} p^x\, q^{n-x}, & for \quad x = 0,1,2,\dots,n, \\ 0, & elsewhere. \end{cases} \tag{3.6}$$

In this distribution n must be positive integer, p must lie in the interval $0 \le p \le 1$ and the variables $x_1, \dots, x_n$ form n Bernoulli trials with parameter p. [Ref. 8: p. 245]

Suppose that P, a random variable, represents the market share of a new brand of a certain product. The value of P is a proportion and can take on any value between 0 and 1. The new brand is considerably different from the other brands of the product, so we are quite uncertain about the share of the market that it will attract. We think that it might attract virtually the entire market for the product (that is, P might be close to 1), it might not be successful at all (that is, P might be close to 0), or it might be moderately successful. Again we assume that P is continuous random variable. We think that low values of P are more likely than high values, and we assess a prior distribution for P that is a special case of a variant of the triangular density function (see Equation 3.4).

In the example of the market share we wish to obtain more information about P. A sample of five consumers of the product is taken: one purchases the new brand and the other four purchase other brands. We also assume that

the process of purchasing this product has n independent Bernoulli trials. That is, the probability that a randomly selected consumer purchases the new brand is equal to P, the market share. The sample information can be represented by the binomial distribution that is as follows :

$$f(X \mid p) = \binom{5}{1} p(1-p)^4.$$

This sampling distribution is graphed in Figure 3.

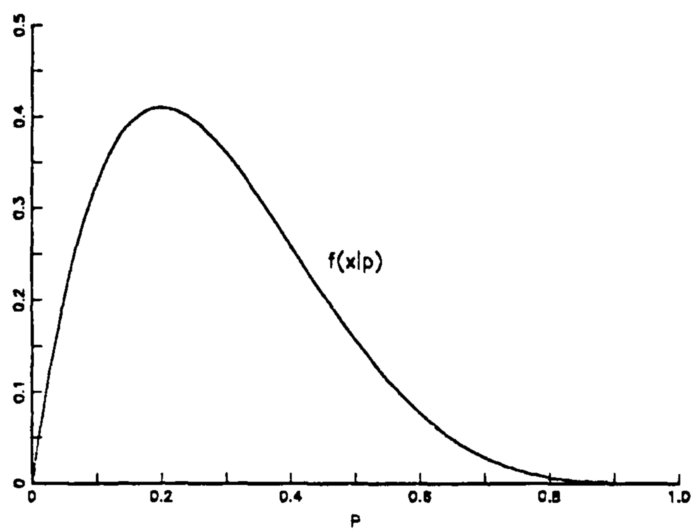

Figure 3. The Sampling Distribution in the Example of the Market Share

Also, the prior distribution of this example is from Equation 3.4, and is illustrated in Figure 4.

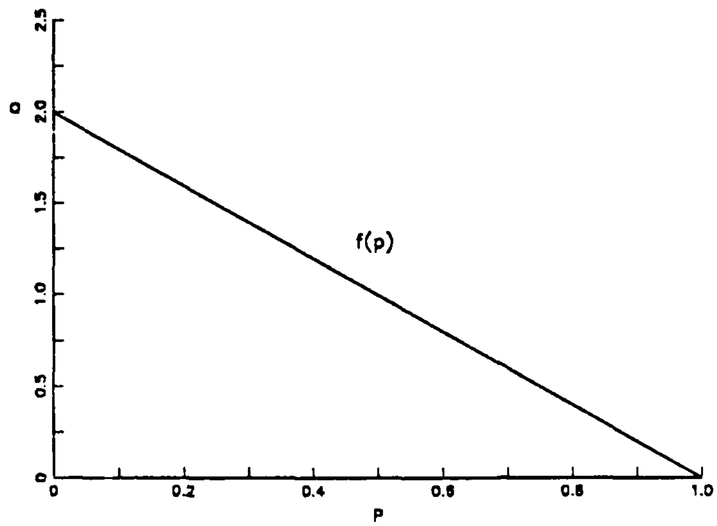

Figure 4. The Prior Distribution in the Example of the Market Share

Applying the version of Bayes' theorem in Equation 3.1, we have the posterior distribution

$$f(p \mid x) = \frac{(2\,(1-p))\ (5p\,(1-p)^4)}{\int_0^1 (2\,(1-p))\ (5p\,(1-p)^4)\,dp}\,.$$

$$= \frac{p\,(1-p)^5}{\int_0^1 p\,(1-p)^5\,dp}\,.$$

The denominator can be stated so that the integral is over a beta density function, yielding

$$\frac{\Gamma(2)\Gamma(6)}{\Gamma(8)} \int_0^1 \frac{\Gamma(8)}{\Gamma(2)\Gamma(6)} p^2(1-p)^6 dp = \frac{1}{42}.$$

Thus , the posterior density function of P is

$$f(p \mid x) = \begin{cases} 42\, p(1-p)^5, & \text{if } 0 \le p \le 1, \\ 0, & \text{elsewhere.} \end{cases}$$

This posterior distribution is illustrated in Figure 5.

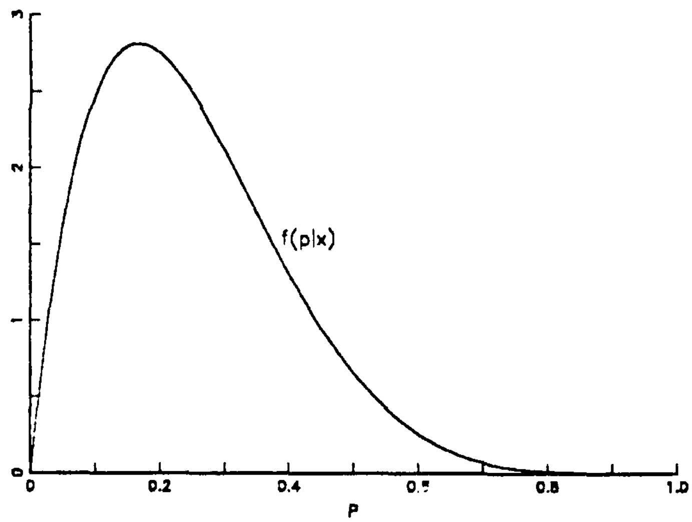

Figure 5. The Posterior Distribution in the Example of the Market Share

This example illustrates how the version of Bayes' theorem in Equation 3.1 provides a convenient way to revise density functions in terms of sample information, as shown by a comparison of Figures 4 and 5.

C. GENERAL DERIVATION OF THE POSTERIOR DISTRIBUTION WHEN THE PRIOR DISTRIBUTION IS TRIANGULAR

Using Bayes' theorem, the prior information (represented by the prior distribution) and the sampling distribution are combined to form the posterior distribution of P. The posterior distribution summarizes our degree of belief of the location of P, given the results of the sample. Of course, the posterior distribution depends on the sampling function as well as on the prior distribution.

At this point, we need to remember the definition of the beta distribution because we will use it in the remainder of this work.

It is said that a random variable P has a beta distribution with parameters α and β ($\alpha > 0$ and $\beta > 0$) if P has a continuous distribution for which the p.d.f. $f(p \mid \alpha, \beta)$ is as follows :

$$f(p \mid \alpha, \beta) = \begin{cases} \dfrac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} p^{\alpha-1}(1-p)^{\beta-1}, & \text{for } 0 < p < 1, \\ 0, & \text{elsewhere}. \end{cases} \tag{3.7}$$

[Ref. 8: p. 294].

In later equations, the density function $f(p \mid \alpha, \beta)$ in Equation 3.7 will be shown as $b(p; \alpha, \beta)$. Having defined the beta distribution for later work, we are ready to seek the posterior distributions that occur with prior triangular distributions and the binomial sampling distribution.

1. Posterior Distribution with Prior Triangular Distributions Having Parameter Pmax

First, we will derive the posterior distribution by using a prior triangular distribution with parameter p_{max} which is going to be more heavily weighted in favor of low values of P rather than high values. Applying the version of Bayes' theorem in Equation 3.1 to combine the prior triangular

distribution in Equation 3.2 and the binomial sampling distribution in Equation 3.6 , we have

$$f_1(p \mid x) = \frac{\binom{n}{x} p^x (1-p)^{n-x} \frac{2}{p_{\max}^2} (p_{\max} - p)}{\int_0^{p_{\max}} \binom{n}{x} p^x (1-p)^{n-x} \frac{2}{p_{\max}^2} (p_{\max} - p) dp},$$

where $0 < p \leq p_{\max} \leq 1$.

If we cancel out some terms, we have

$$f_1(p \mid x) = \frac{p^x (1-p)^{n-x} (p_{\max} - p)}{\int_0^{p_{\max}} p^x (1-p)^{n-x} (p_{\max} - p) dp}.$$

If we multiply terms in the denominator, our posterior becomes

$$f_1(p \mid x) = \frac{p^x (1-p)^{n-x} (p_{\max} - p)}{p_{\max} \int_0^{p_{\max}} p^x (1-p)^{n-x} dp - \int_0^{p_{\max}} p^{x+1} (1-p)^{n-x} dp}.$$

We multiply the denominator by the same terms to create the beta density functions, $b(x+1,n-x+1)$, under the first integral, and $b(x+2,n-x+1)$ under the second integral. Then the denominator is

$$\left[\left(p_{\max} \frac{\Gamma(x+1)\Gamma(n-x+1)}{\Gamma(n+2)}\right) \int_0^{p_{\max}} b(p;x+1,n-x+1) dp\right] -$$

$$\left[\left(\frac{\Gamma(x+2)\Gamma(n-x+1)}{\Gamma(n+3)}\right) \int_0^{p_{\max}} b(p;x+2,n-x+1) dp\right].$$

By using the property of the gamma function that

$$\Gamma(\alpha) = (\alpha - 1)\,\Gamma(\alpha - 1), \quad (\textit{if} \quad \alpha > 1),$$

the denominator becomes

$$\left[\left(p_{\max}\frac{\Gamma(x+1)\Gamma(n-x+1)}{\Gamma(n+2)}\right)\int_0^{p_{\max}} b(p;x+1,n-x+1)dp\right] -$$

$$\left[\left(\frac{(x+1)\Gamma(x+1)\Gamma(n-x+1)}{n+2\Gamma(n+2)}\right)\int_0^{p_{\max}} b(p;x+2,n-x+1)dp\right].$$

If we combine terms, we have the posterior distribution $f_1(p \mid x)$ as

$$\frac{\left(\frac{\Gamma(n+2)}{\Gamma(x+1)\Gamma(n-x+1)}\right)\left[(p^{x}(1-p)^{n-x}p_{\max}) - (p^{x+1}(1-p)^{n-x})\right]}{\left[p_{\max}\int_0^{p_{\max}} b(p;x+1,n-x+1)dp\right] - \left[\frac{x+1}{n+2}\int_0^{p_{\max}} b(p;x+2,n-x+1)dp\right]}.$$

When we multiply the terms in the numerator, multiply the second term by $\left(\frac{x+1}{n+2} \times \frac{n+2}{x+1}\right)$, and then use the property of the gamma function that $\Gamma\ (n+3) = (n+2)\ \Gamma\ (n+2)$ we will obtain two forms of the beta density function with parameters

- $\alpha_1 = x + 1$,
- $\alpha_2 = x + 2$,
- *and* $\beta_1 = \beta_2 = n - x + 1$.

The posterior distribution, $f_1(p \mid x)$, becomes

$$\frac{[p_{\max} b(p;x+1,n-x+1)] - \left[\frac{x+1}{n+2} b(p;x+2,n-x+1)\right]}{\left[p_{\max}\int_0^{p_{\max}} b(p;x+1,n-x+1)dp\right] - \left[\frac{x+1}{n+2}\int_0^{p_{\max}} b(p;x+2,n-x+1)dp\right]}.$$

We notice that the denominator has two forms of the beta cumulative distribution function with the same parameters α_1, α_2, β_1 and β_2 , which is shown as $B(p_{max}; \alpha, \beta)$. Therefore we have

$$\frac{[p_{max} b(p\,;x+1, n-x+1)] - \left[\frac{x+1}{n+2} b(p\,;x+2, n-x+1)\right]}{[p_{max} B(p_{max}\,;x+1, n-x+1)] - \left[\frac{x+1}{n+2} B(p_{max}\,;x+2, n-x+1)\right]}.$$

So finally, our posterior distribution becomes

$$f_1(p \mid x) = \frac{[p_{max} b(p\,;\alpha_1, \beta_1)] - \left[\frac{x+1}{n+2} b(p\,;\alpha_2, \beta_2)\right]}{[p_{max} B(p_{max}\,;\alpha_1, \beta_1)] - \left[\frac{x+1}{n+2} B(p_{max}\,;\alpha_2, \beta_2)\right]}, \quad (3.8)$$

where

- p is the probability of success ($0 < p \le p_{max} \le 1$).
- x is the number of successes that occurred in n independent Bernoulli trials.
- B $(p_{max}; \alpha_1, \beta_1)$ and B $(p_{max}; \alpha_2, \beta_2)$ show the beta c.d.f.,
- b $(p; \alpha_1, \beta_1)$ and b $(p; \alpha_2, \beta_2)$ show the beta p.d.f.,
- $\alpha_1 = x + 1$,
- $\alpha_2 = x + 2$,
- and $\beta_1 = \beta_2 = n - x + 1$.

Using a prior triangular distribution with parameter p_{max} and a binomial sampling distribution, we have the posterior distribution in Equation 3.8. This posterior distribution shows that the size of the confidence interval for the proportion P depends upon the parameter of the prior (p_{max}), the sample size n, and the number of successes x .

We remember from the classical method that we need to know the number x of successes (before sampling) to determine the number n of samples. So, prior to sampling, we will make an assumption about x that it is

equal to its expected value which is the mean of the prior triangular distribution with parameter p_{max}, multiplied by the number of samples, or

$$x = \left(\frac{p_{\max}}{3} n \right).$$

Applying the result of this assumption into Equation 3.8, we have the posterior distribution as

$$\frac{\left[p_{\max} b\left(\frac{np_{\max}+3}{3}, \frac{3n - np_{\max}+3}{3} \right) \right] - \left[\frac{np_{\max}+3}{3n+6} b\left(\frac{np_{\max}+6}{3}, \frac{3n - np_{\max}+3}{3} \right) \right]}{\left[p_{\max} B\left(\frac{np_{\max}+3}{3}, \frac{3n - np_{\max}+3}{3} \right) \right] - \left[\frac{np_{\max}+3}{3n+6} B\left(\frac{np_{\max}+6}{3}, \frac{3n - np_{\max}+3}{3} \right) \right]}.$$

If we let

- $\alpha_1^* = \dfrac{n\, p_{\max} + 3}{3}$,
- $\alpha_2^* = \dfrac{n\, p_{\max} + 6}{3}$,
- $\beta_1^* = \beta_2^* = \dfrac{3n - n\, p_{\max} + 3}{3}$,
- and $N_1 = \dfrac{n\, p_{\max} + 3}{3n + 6}$.

one of the our posterior distributions becomes

$$f_1(p \mid x) = \frac{\left[p_{\max}\, b(p : \alpha_1^*, \beta_1^*) \right] - \left[N_1\, b(p : \alpha_2^*, \beta_2^*) \right]}{\left[p_{\max}\, B(p_{\max} ; \alpha_1^*, \beta_1^*) \right] - \left[N_1\, B(p_{max} : \alpha_2^*, \beta_2^*) \right]}, \tag{3.9}$$

where $\alpha_1^*, \alpha_2^*, \beta_1^*$, and $\beta_2^* > 0$.

The final form of one of our posterior distributions continues, of course, to have two forms of the beta density function with parameters $\alpha_1^*, \alpha_2^*, \beta_1^*, \beta_2^*$ in the numerator and two forms of the cumulative distribution function of the beta distribution with same parameters $\alpha_1^*, \alpha_2^*, \beta_1^*, \beta_2^*$ in the denominator. This means that existing computer programs for both the beta density function and

the beta c.d.f. can be employed in our computationial work to relate desired confidence interval size to the sample size.

2. Posterior Distribution with Prior Triangular Distributions with Parameter Pmin

Now we will derive the posterior distribution by using a prior triangular distribution with parameter p_{min}. Here, $f(p)$ is going to be more heavily weighted in favor of high values of P rather than low values. Applying Bayes' theorem from Equation 3.1 to combine the prior triangular distribution in Equation 3.3 and the binomial sampling distribution in Equation 3.6,

$$f_2(p \mid x) = \frac{\binom{n}{x} p^x (1-p)^{n-x} \dfrac{-2(p_{\min} - p)}{(p_{\min} - 1)^2}}{\displaystyle\int_{p_{\min}}^{1} \binom{n}{x} p^x (1-p)^{n-x} \frac{-2(p_{\min} - p)}{(p_{\min} - 1)^2}\, dp},$$

where $0 \le p_{min} \le p < 1$. After we apply the same steps shown above in the derivation of the $f_1(p \mid x)$, the posterior distribution, $f_2(p \mid x)$, becomes

$$\frac{[p_{\min} b(p ; \alpha_3, \beta_3)] - \left[\dfrac{x+1}{n+2}\, b(p ; \alpha_4, \beta_4) \right]}{\{p_{\min}[1 - B(p_{\min} ; \alpha_3, \beta_3)]\} - \left\{ \dfrac{x+1}{n+2}\, [1 - B(p_{\min} ; \alpha_4, \beta_4)] \right\}}. \tag{3.10}$$

Here

- p is the probability of success ($0 \le p_{min} \le p < 1$).
- x is the number of successes that occured in n independent Bernoulli trials,
- B (p_{min}; α_3, β_3) and B (p_{min}; α_4, β_4) show the beta c.d.f.,
- b (p; α_3, β_3) and b (p; α_4, β_4) show the beta p.d.f.,
- $\alpha_3 = x + 1$,
- $\alpha_4 = x + 2$,
- and $\beta_3 = \beta_4 = n - x + 1$.

Using a prior triangular distribution with parameter $p_{\min}$ and a binomial sampling distribution, we have the posterior distribution in Equation 3.10. This posterior distribution shows that the size of the confidence interval depends upon the parameter of the prior ($p_{\min}$), the sample size n, and the number of successes x.

At this point, we also need to know the number x of successes (before sampling) to determine the number n of samples. So we will make an assumption about x that it is equal to the mean of the prior triangular distribution with parameter $p_{\min}$ multiplied by the number of samples, or

$$x = \left(\frac{n}{3}\,(p_{\min} + 2)\right).$$

Applying the result of this assumption to the Equation 3.10, we now have the posterior distribution $f_2(p \mid x)$ as

$$\frac{\left[p_{\min} b\left(\frac{np_{\min}+2n+3}{3}, \frac{n-np_{\min}+3}{3}\right)\right] - \left[\frac{np_{\min}+2n+3}{3n+6}\, b\left(\frac{np_{\min}+2n+6}{3}, \frac{n-np_{\min}+3}{3}\right)\right]}{p_{\min}\left[1 - B\left(\frac{np_{\min}+2n+3}{3}, n - n\frac{p_{\min}-3}{3}\right)\right] - \frac{np_{\min}+2n+3}{3n+6}\left[1 - B\left(\frac{np_{\min}+2n+6}{3}, \frac{n-np_{\min}+3}{3}\right)\right]}.$$

If we let

- $\alpha_3^{*} = \dfrac{n\,p_{\min} + 2n + 3}{3}$,
- $\alpha_4^{*} = \dfrac{n\,p_{\min} + 2n + 6}{3}$,
- $\beta_3^{*} = \beta_4^{*} = \dfrac{n - n\,p_{\min} + 3}{3}$,
- and $N_2 = \dfrac{n\,p_{\min} + 2n + 3}{3n + 6}$,

then our second posterior distribution becomes

$$f_2(p \mid x) = \frac{\left[p_{\min} b(p : \alpha_3^{*}, \beta_3^{*})\right] - \left[N_2 b(p : \alpha_4^{*}, \beta_4^{*})\right]}{\left\{p_{\min}\left[1 - B(p_{\min} ; \alpha_3^{*}, \beta_3^{*})\right]\right\} - \left\{N_2\left[1 - B(p_{min} ; \alpha_4^{*}, \beta_4^{*})\right]\right\}}, \quad (3.11)$$

where $\alpha_3^{*}, \alpha_4^{*}, \beta_3^{*}$, and $\beta_4^{*} > 0$.

The final form of our second posterior distribution has also two forms of the beta density function with the paran.ɔters $\alpha_3^{*}, \alpha_4^{*}, \beta_3^{*}, \beta_4^{*}$ in the numerator and two forms of the beta cumulative distribution function with the same parameters $\alpha_3^{*}, \alpha_4^{*}, \beta_3^{*}, \beta_4^{*}$ in the denominator.

We will use Equation 3.9 and Equation 3.11 as our posterior distributions in computer programs which were written in APL.

In the next chapter we will present computer programs, tables, and graphs that can be used by a decision maker to determine the required sample size to obtain a desired size for a 95% Bayesian confidence interval for a probability or proportion. Also, we will show that in the Bayesian method, the desired sample size may be smaller than the sample sizes obtained by using the classical methods of Chapter II.

IV. DETERMINING THE DESIRED SAMPLE SIZE TO ESTIMATE PROPORTIONS USING THE BAYESIAN METHOD

In this chapter we will discuss the Bayesian method by making use of subjective probabilities measuring degrees of belief in order to determine the number of samples for proportions. As we mentioned in Chapter III, these probabilities are called the prior distribution. Thus when using the Bayesian method, this prior distribution summarizes the decision maker's subjective degree of belief about the unknown values of proportions. When the decision maker has this subjective prior information (degree of belief) about bounds for the unknown proportion which can be described by prior triangular distributions having parameters p_{max} or p_{min}, the sample size can be determined by using these triangular distributions, the desired confidence level, and the confidence interval size.

First, we will discuss some of the considerations the decision maker might make which would lead to the selection of a prior triangular distribution. In the next section, after developing the relationship between the desired sample size and the decision maker's prior information represented by one of the two variants of the triangular distribution, we will explain the tables which are related to the Bayesian interval sizes and the number of samples. Next, we will explain the graphs to determine the number of samples for proportions. Finally, we will discuss the Bayesian method using the prior triangular distribution in which the sample sizes may be smaller than the sample sizes obtained by using the classical method in Table 1 on Page 8. Throughout this chapter we will also explain how the computer programs work in determining the sample size to estimate proportions.

A. DETERMINING THE PRIOR TRIANGULAR DISTRIBUTIONS AND THEIR PARAMETERS

Under what conditions might one of the two forms of the triangular density function be reasonable as priors? We can use the forms of the triangular

distribution as priors to represent skewing without having extensive additional information about the prior distribution. When the decision maker feels skewing is present, a triangular prior is an improvement over the prior uniform distribution, and the prior triangular distribution is also less complicated than a prior beta distribution where two parameters, which may be subjective, must be stated to fit the decision maker's prior information. For a prior triangular distribution, for example, the only information needed from the decision maker is that low values of P are more likely than high values (e.g., P = proportion nonconforming) and a statement of p_{max}, which could be 1.0. Alternately, the decision maker may feel that high values of P are more likely, in which case he or she needs only a value for p_{min}, which could be 0.

First, we consider a random variable P with the triangular density function of the form in Equation 3.2 which is as follows:

$$f_1(p) = \begin{cases} \dfrac{2}{p_{\max}^2} \; (p_{\max} - p), & for \;\; 0 < p \le p_{\max} \le 1. \\ 0, & elsewhere. \end{cases}$$

This prior triangular distribution is going to be more heavily weighted in favor of low values of P rather than high values. One way the decision maker can decide the parameter of this prior triangular distribution is by using the mean of this prior triangular distribution, which is $E(P) = \frac{p_{max}}{3}$ and must be selected between 0 and 0.33333. For example, if the decison maker guesses the mean value as 0.2, the parameter can be found as follows:

$$p_{\max} = 3\,E(\,P\,) = 0.6.$$

At the same time, the value of parameter p_{max} reflects the amount of positive skewing.

Next, we consider the other form of the triangular density function in Equation 3.3 which is as follows:

$$f_2(p) = \begin{cases} -\dfrac{2\ (p_{\min} - p)}{(p_{\min} - 1)^2}, & for\ \ 0 \le p_{\min} \le p < 1, \\ 0, & elsewhere. \end{cases}$$

This form of the triangular density function with parameter $p_{\min}$ is going to be more heavily weighted in favor of high values of proportions rather than low values. Again, one way by which the decision maker can decide the parameter of this prior triangular distribution is to use the mean of the prior triangular distribution, which is $E(P) = \frac{1}{3}(p_{\min} + 2)$ and must be selected between 0.66666 and 1. For example, if the decison maker guesses the mean value E(P) as 0.8, the parameter can be found as follows:

$$p_{\min} = 3E(P) - 2 = 0.4.$$

At the same time, the value of parameter $p_{\min}$ reflects the amount of negative skewing.

These examples show how triangular distributions allow the decision maker to express prior beliefs about the values of P when only limited information is available. We will next explain how we find the Bayesian bounds and interval sizes.

B. FINDING THE BAYESIAN INTERVAL SIZES AND BOUNDS

We gave the derivation of the posterior distributions in Chapter III when the priors are in the form of the triangular density function. Results from these posterior distributions are as follows:

$$f_1(p \mid x) = \frac{\left[p_{\max}\, b(p\,;\alpha_1^{*}, \beta_1^{*})\right] - \left[N_1\, b(p\,;\alpha_2^{*}, \beta_2^{*})\right]}{\left[p_{\max}\, B(p_{\max}\,;\alpha_1^{*}, \beta_1^{*})\right] - \left[N_1\, B(p_{max}\,;\alpha_2^{*}, \beta_2^{*})\right]}$$

where $\alpha_1^{*}, \alpha_2^{*}, \beta_1^{*}$, and $\beta_2^{*} > 0$, and

$$f_2(p \mid x) = \frac{\left[p_{\min}\, b(p\,;\alpha_3^{*}, \beta_3^{*})\right] - \left[N_2\, b(p\,;\alpha_4^{*}, \beta_4^{*})\right]}{\left\{p_{\min}\left[1 - B(p_{\min}\,;\alpha_3^{*}, \beta_3^{*})\right]\right\} - \left\{N_2\left[1 - B(p_{min}\,;\alpha_4^{*}, \beta_4^{*})\right]\right\}}.$$

where α_3^*, α_4^*, β_3^*, and $\beta_4^* > 0$. In these expressions we assumed that the number of successes is equal to the mean of the prior distribution times n. The posterior distributions in the above forms are here for the purpose of finding sample size. As we mentioned in Chapter III, they are in the form of linear combinations of two beta density functions in the numerator, and the linear combination of two beta cumulative distribution functions in the denominator. The parameters are functions of the sample size n, and p_{max} or p_{min}. The denominators are not functions of the random variable P.

If we specify a value for the sample size n and a value for p_{min} or p_{max}, we can compute (for, say, a 95 percent confidence level) the cumulative distribution functions at 0.025 and 0.975 for the posterior density functions. We notice that the term in the denominator is constant for given p_{max} and n.

For example, the cumulative distribution functions of the posterior distribution $f_1(p \mid x = \frac{p_{max}}{3})$ at 0.025 and 0.975 are

$$F_1(p.lo) = \frac{p_{max}\int_0^{p.lo} b(p)dp \; - \; N_1\int_0^{p.lo} b(p)dp}{p_{max}\int_0^{p_{max}} b(p;\, \alpha_1^*,\, \beta_1^*)dp - N_1\int_0^{p_{max}} b(p;\, \alpha_2^*,\, \beta_2^*)dp} = 0.025. \tag{4.1}$$

and

$$F_1(p.up) = \frac{p_{max}\int_0^{p.up} b(p)dp - N_1\int_0^{p.up} b(p)dp}{p_{max}\int_0^{p_{max}} b(p;\, \alpha_1^*,\, \beta_1^*)dp - N_1\int_0^{p_{max}} b(p;\, \alpha_2^*,\, \beta_2^*)dp} = 0.975. \tag{4.2}$$

Using the above equations we can obtain the lower and upper bounds of the 95 percent Bayesian confidence interval that would result had the number of successes in the sample been equal to the mean number from the prior distribution. Finally, we will find the interval sizes by subtracting the upper bound from the lower bound.

For the above procedure, we developed the APL programs named PMINIMUM and PMAXIMUM, which are given in Appendix B and C, respectively. The program, PMINIMUM, is for the posterior distribution with the prior triangular density function having parameter p_{min}, and PMAXIMUM is for the other posterior distribution with the prior triangular distribution having parameter p_{max}. Both programs are the main programs used in our analysis. Each program is interactive and the user is required to enter the bound of the prior triangular distribution (p_{max} *or* p_{min}) and the sample sizes.

For example, the APL program, PMAXIMUM, computes the upper bounds, the lower bounds, and the Bayesian interval sizes using Equations 4.1 and 4.2. Also it uses the APL program BETA which was designed at the Naval Postgraduate School to compute the beta density function. This program is given in Appendix D. It should be noticed that if the total parameter value of any of the beta distributions in the posterior distributions exceeds 255 (i.e., $\alpha_3^* + \beta_3^* \geq 255$ *or* $\alpha_4^* + \beta_4^* \geq 255$), BETA cannot compute the beta density function.

In the next section, we will explain how the decion maker can use these tables.

C. DETERMINING THE SAMPLE SIZES WITH TABLES

Decision makers can use tables to facilitate their determination of the sample size using the prior triangular distribution and a desired confidence level.

Let us explain with examples how the decision maker can use these tables. Suppose that the decision maker's prior triangular distribution parameters are $p_{max} = 1$ or $p_{min} = 0$. Also suppose that the decision maker desires the Bayesian interval size (2A) to be 0.20 for estimating the proportion, with a 95 percent confidence level. The APL programs create tables similar to Tables 2 and 3 using the parameters $p_{max} = 1$ or $p_{min} = 0$, respectively. Then the decision maker can find the sample size, 81, from Table 2 or 3. This is the desired sample size n that reflects both the decision maker's subjective bounds and a 95 percent confidence level.

As we mentioned in Chapter III, the prior triangular distributions having parameter $p_{max} = 1$ and $p_{min} = 0$ are symmetric which also holds for $p_{max} = 0.8$ and $p_{min} = 0.2$ or other situations. Therefore, we notice that their Bayesian interval sizes are the same (see Table 2 and 3), but also we should notice that they have different lower and upper bounds with the same Bayesian interval size.

Table 2. SAMPLE SIZES AND BAYESIAN INTERVALS USING THE TRIANGULAR PRIOR DISTRIBUTION WITH PARAMETER PMAX=1.0

Sample Size n	Lower Bound	Upper Bound	Bayesian Interval Size = 2A
1	0.0259	0.7890	0.7631
2	0.0397	0.7484	0.7087
3	0.0527	0.7164	0.6637
4	0.0647	0.6905	0.6258
5	0.0755	0.6690	0.5935
6	0.0852	0.6509	0.5656
7	0.0940	0.6353	0.5412
8	0.1020	0.6217	0.5197
9	0.1093	0.6097	0.5005
10	0.1159	0.5991	0.4832
20	0.1607	0.5334	0.3728
30	0.1857	0.5001	0.3144
40	0.2022	0.4791	0.2769
50	0.2142	0.4644	0.2502
60	0.2233	0.4534	0.2300
70	0.2306	0.4447	0.2140
80	0.2366	0.4376	0.2010
81	0.2372	0.4370	0.1998
82	0.2377	0.4364	0.1987
83	0.2382	0.4358	0.1975
84	0.2388	0.4352	0.1964
85	0.2393	0.4346	0.1953
90	0.2417	0.4318	0.1901
100	0.2460	0.4268	0.1807
110	0.2498	0.4225	0.1727
120	0.2531	0.4187	0.1656
130	0.2560	0.4154	0.1593
140	0.2587	0.4124	0.1537
150	0.2611	0.4097	0.1487
160	0.2632	0.4073	0.1441

Table 3. SAMPLE SIZES AND BAYESIAN INTERVALS USING THE TRIANGULAR PRIOR DISTRIBUTION WITH PARAMETER PMIN = 0.0

Sample Size n	Lower Bound	Upper Bound	Bayesian Interval Size = 2A
1	0.2110	0.9741	0.7631
2	0.2516	0.9603	0.7087
3	0.2836	0.9473	0.6637
4	0.3095	0.9353	0.6258
5	0.3310	0.9245	0.5935
6	0.3491	0.9148	0.5656
7	0.3647	0.9060	0.5412
8	0.3783	0.8980	0.5197
9	0.3903	0.8907	0.5005
10	0.4009	0.8841	0.4832
20	0.4666	0.8393	0.3728
30	0.4999	0.8143	0.3144
40	0.5209	0.7978	0.2769
50	0.5356	0.7858	0.2502
60	0.5466	0.7767	0.2300
70	0.5553	0.7694	0.2140
80	0.5624	0.7634	0.2010
81	0.5630	0.7628	0.1998
82	0.5636	0.7623	0.1987
83	0.5642	0.7618	0.1975
84	0.5648	0.7612	0.1964
85	0.5654	0.7607	0.1953
90	0.5682	0.7583	0.1901
100	0.5732	0.7540	0.1807
110	0.5775	0.7502	0.1727
120	0.5813	0.7469	0.1656
130	0.5846	0.7440	0.1593
140	0.5876	0.7413	0.1537
150	0.5903	0.7389	0.1487
160	0.5927	0.7368	0.1441

It is true that

- Upper Bound (in Table 2) = 1 - Lower Bound (in Table 3) or
- Lower Bound (in Table 2) = 1 - Upper Bound (in Table 3).

Therefore after obtaining the tables with the triangular density function using p_{min}, it is easy to obtain other tables with the triangular density function using p_{max}.

Tables such as Table 2 and Table 3 with various values of p_{max} and p_{min} that can be used to determine the sample size needed to produce a desired confidence level to estimate a proportion are located in Appendices E and F. In the next section, we will explain how the decision maker can use graphs to determine the sample size to estimate proportions.

D. DETERMINING THE SAMPLE SIZES WITH GRAPHS

In this section, we will provide graphs to assist in determining the number of samples by using triangular distributions with various parameters as priors.

Programs PMINIMUM and PMAXIMUM create vectors in the APL workspace of the lower bounds, the upper bounds, and Bayesian interval sizes. If we plot the vector of the sample sizes versus the vector of the lower bounds and the upper bounds, we can obtain the graph illustrated in Figure 6. From Figure 6, it can be seen that when the sample sizes increase, the Bayesian interval sizes decrease.

If we plot the vector of the sample sizes versus the size of the 95 percent Bayesian interval we obtain the graph illustrated in Figure 7. Let us explain with an example how the decision maker can use this graph. Suppose that the decision maker's prior triangular distribution parameter is $p_{min}=0$. In addition, suppose that the decision maker desires the Bayesian interval size to be 0.20 with a 95 percent confidence. First, the decision maker, using Figure 7, must find 0.20 on the ordinate and then move across the graph to where 0.20 intercepts the curve. The decision maker can then read the sample size, approximately 81, on the abscissa.

The decision maker can find graphs in Appendix G with various parameters of the prior triangular distribution to determine the sample sizes

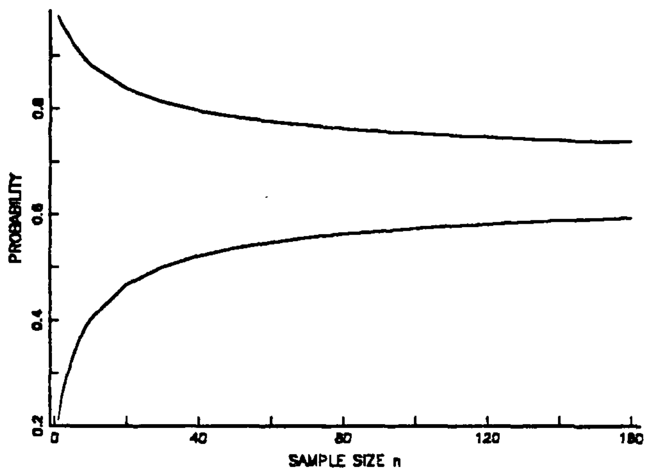

Figure 6. Number of Samples vs the Bounds of the 95% Bayesian Interval with a Triangular Prior Distribution Having Parameter Pmin = 0

needed to obtain a desired 95 percent confidence level for proportions. In the next section we will study the sensitivity of sample size to the bounds in the prior distribution.

E. SENSITIVITY OF SAMPLE SIZE TO THE PARAMETERS IN THE PRIOR DISTRIBUTION

At this point the decision maker may be interested in this question: how will variations in the prior distribution affect the ultimate decision? Or, in other words, how sensitive is the sample size n to the (possibly) guessed value of the bound (p_{max} or p_{min}) in the prior distribution? To answer these questions we will change p_{max} or p_{min} with the sample size n held constant, and we will look at the values of the Bayesian interval size 2A. For example, if we want the Bayesian interval size 2A to be 0.2 and guess $p_{min} = 0.2$, we obtain n = 70. Now

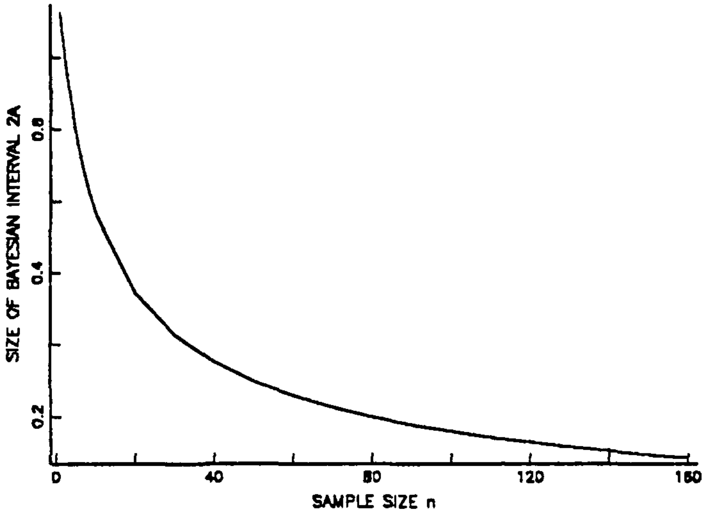

Figure 7. Number of Samples vs the Size of the 95% Bayesian Interval with a Triangular Prior Distribution Having Parameter Pmin = 0

if we change our guess, e.g., $p_{min} = 0.1$, and use $n = 70$, we obtain $2A = 0.2079$. Erroring in the other direction, for $p_{min} = 0.3$ and $n = 70$, we obtain $2A = 0.195$. Thus for this example, the sample size n and the choice of bounds p_{max} or p_{min} appear to be relatively insensitive. When the sample sizes are bigger, the sample size n is reasonably insensitive to the choice of guessed values of p_{min} or p_{max} (see Figure 8).

Next, we will make a comparison between methods.

F. COMPARISON OF THE CLASSICAL METHOD AND THE BAYESIAN METHOD USING THE TRIANGULAR PRIOR DISTRIBUTION

Let us compare the results obtained by using our approach versus the classical method. The classical method for determining the desired sample

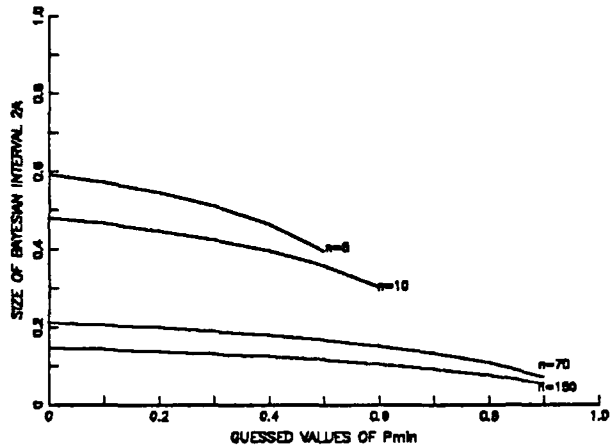

Figure 8. The Sensitivity of C.I. Size to the Guessed Value of Pmin

sizes to estimate proportions requires the decision maker's guess about the probability of success and the length of the confidence interval (2A). The Bayesian method, our alternative approach, needs 2A and the bounds or parameters of the prior density function and direction of skewing. Also, in this discussion we note that

> For any finite sample size, the Bayesian estimate is 'shaded' toward the prior mean, the best guess for P before any sample values were taken. This effect disappears as n increases indefinitely.[Ref. 1: p. 566]

Suppose that the bound of the prior triangular distribution is $p_{min} = 0$ and the desired interval size (2A) is 0.20. Then we find E(P) = 0.6666 as the mean of this prior. Suppose that this mean value from the Bayesian method is equal to the decision maker's guess about the probability of success in the classical method. By using this value in Table 1 on page 8, the classical method requires 86 as the sample size. At the same time, if we use the tables based

on the Bayesian method we find the sample size to be 81 from Table 3. Other examples are shown in Table 4.

So, if we compare the results of these two methods which meet the above requirements, we realize that the results of the Bayesian method are quite favorable to those obtained using the classical method. When the values of the sample sizes are smaller, the values of the sample sizes based on Bayesian method are quite different. This is apparent from the results in Table 4. Larson says, "The difference between the Bayesian values, and the classical approach, disappears as n increases" [Ref. 1: p. 573]. Also, we realize that when the sample sizes become larger, the posterior distribution becomes less dependent on the subjective prior information and more dependent on the objective sample information.

Table 4. COMPARISON OF THE CLASSICAL AND THE BAYESIAN METHODS

Prior Bound p_{min}	Prior E(P)	Size of Interval 2A	Sample Size from Bayesian Method	Sample Size from Classical Method
0	0.66	0.15	147	152
0	0.66	0.20	81	86
0	0.66	0.25	50	55
0	0.66	0.30	33	38
0.1	0.7	0.15	139	144
0.1	0.7	0.20	76	81
0.1	0.7	0.25	47	52
0.1	0.7	0.30	31	36
0.2	0.73	0.15	129	135
0.2	0.73	0.20	70	76
0.2	0.73	0.25	43	49
0.2	0.73	0.30	29	34

In the next chapter, we will summarize our study, and we will give some suggestions for further research and study.

V. SUMMARY AND SUGGESTIONS FOR FURTHER RESEARCH AND STUDY

Determination of the number of items to be tested remains an important problem in military operational test and evaluation, particularly when a proportion, such as an item's reliability, is to be estimated. One well known approach is to state a requirement for the size of the confidence interval that will estimate the proportion, and then use that requirement to determine the sample size.

Application of Bayesian statistics (which employ prior information about the proportion to be estimated) can reduce the number of observations needed. Floropoulos [Ref. 3] studied the case where, prior to sampling, the decision maker might be able to bound the proportion, but was uncertain about it otherwise. Manion [Ref. 2] examined the sample size determination question when enough prior information was available to specify a beta distribution as a prior. The study in this thesis looked at the case where there was more information present than the uncertainty represented by a uniform distribution, but not enough detail to use a beta prior.

In this chapter, we will summarize how we used the Bayesian method with the triangular priors to obtain the sample sizes to estimate proportions. Finally, we will give some suggestions for additional studies.

A. SUMMARY

Throughout this study, we described the Bayesian method to determine the desired sample size that is needed to estimate proportions with a $(1-\alpha)$ 100 confidence when a prior distribution is given to a proportion.

First, we described a classical method to determine the sample size for estimating proportions using confidence intervals. Then, we described an alternative to this approach which was the Bayesian method. We studied the three parts of this Bayesian method: the prior distribution, the sampling distribution, and the posterior distribution. When using the Bayesian method, the prior distribution expresses the decision maker's degree of belief of the

location of proportion P prior to sampling, and the posterior distribution expresses the decision maker's degree of belief of the location of proportion P given the results of the sample. We developed and used the two forms of the triangular density function as our priors. Using Bayes' theorem, we combined these two prior triangular distributions and the binomial sampling distribution to form the posterior distributions. The forms of these posterior distributions had two forms of the beta density function in the numerator and two forms of the beta cumulative distribution function in the denominator. Then using these two posterior distributions, we developed computer programs, tables, and graphs that can be used by a decision maker to determine the desired sample size to obtain a 95 percent confidence level to estimate proportions using Bayesian intervals. We also explained how the decision maker might select the prior triangular distributions and their bounds, and how decision makers can use tables and graphs to facilitate their determination of the sample size in some decision making applications.

Finally, we showed that results from the Bayesian method are quite favorable to those obtained using the classical method. When the values of the sample size are small, the values of the sample sizes based on the Bayesian method are quite an improvement. We also showed that when the values of the sample size are large, the sample size n is reasonably insensitive to the choice of bounds, p_{max} or p_{min}.

In the next section we will suggest some additional studies.

B. SUGGESTIONS FOR FURTHER RESEARCH AND STUDY

In previous studies, the prior beta distribution allowed better control of the representation of the decision maker's prior beliefs, while the prior uniform distribution did not provide a great deal of flexibility. In our study, prior triangular distributions did not provide exceptional flexibility in selecting priors but we realized that the use of prior triangular distributions is less complicated than using the prior beta distribution. In other words, these two prior triangular distributions can be used when estimations about proportions are made about the minimum or maximum values of the random variable P

and the skewing of the prior. At this point we suggest that when estimations are made about the modal values of the random variable, another prior density function could have a different triangular shape with the following p.d.f.,

$$f(x) = \begin{cases} \dfrac{2(x-a)}{(b-a)(c-a)}, & a \leq x \leq b, \\ \dfrac{2(c-x)}{(c-b)(c-a)}, & b < x \leq c, \\ 0, & elsewhere, \end{cases}$$

where $a \leq b \leq c$. Here, note that both upper and lower bounds can be changed at the same time.

Also, in addition to 95 percent confidence, all tables and graphs could be developed using other confidence levels, such as 90%, 97.5%, and 99%. An additional study which could be made would determine the number of samples for estimating proportions if nonparametric methods are to be used.

It is hoped that the work presented here will be useful to experimenters, decision makers, and test planners in deciding how big a sample, or how many trials must be done, in order to estimate a proportion or a probability.

APPENDIX A. DERIVATION OF THE TWO VARIANTS OF THE TRIANGULAR DENSITY FUNCTION

We will use basic calculus to derive the two variants of the triangular density function. Both density functions are linear functions, one with negative slope for random variable P defined $0 < p \leq p_{max} \leq 1$, and the other with positive slope for $0 \leq p_{min} \leq p < 1$. First, we will give the definition of slope that is as follows:

If $P_1 = (x_1, y_1)$ and $P_2 = (x_2, y_2)$ are points on a nonvertical line ℓ, the slope of ℓ is defined by the ratio [Ref. 9: p. 19]

$$m = \frac{y_2 - y_1}{x_2 - x_1}. \tag{A.1}$$

Then we will define the equation of ℓ. Suppose ℓ is a line with slope m which contains the point (x_1, y_1). To find the equation for ℓ we let p=(x,y) be an arbitrary point on ℓ. Then, by Equation A.1, we obtain

$$m = \frac{y - y_1}{x - x_1},$$

so

$$y - y_1 = m(x - x_1). \tag{A.2}$$

Now we can find the equation of the first variant of the triangular distribution with parameter p_{max} for the line through (0,a) and (p_{max} ,0) (see Figure 1). First, we will find the slope by using Equation A.1, that is

$$m = \frac{0 - a}{p_{max} - 0} = -\frac{a}{p_{max}},$$

where $0 < p \leq p_{max} \leq 1$ and $2 \leq a < \infty$. Using the point (0,a) and the slope $m = -\frac{a}{p_{max}}$ in Equation A.2 then gives

$$y - a = -\frac{a}{p_{\max}}(p - 0).$$

Setting $y = f(p)$, we then have

$$f(p) = \frac{a}{p_{\max}}(p_{\max} - p). \tag{A.3}$$

We will find an equation for "a" by using the following property of density functions,

$$\int_0^{p_{\max}} \frac{a}{p_{\max}}(p_{\max} - p)dp = 1,$$

which implies that

$$a = \frac{2}{p_{\max}}.$$

Substituting "a" into Equation A.3, we have the triangular density function as

$$f(p) = \frac{2}{p_{\max}^2}(p_{max} - p), \tag{A.4}$$

where $0 < p \leq p_{\max} \leq 1$.

Also, we need to remember the definition of the expectation for a continuous distribution to find the means of the two variants of the triangular density function. If a random variable P has a continuous distribution for which the p.d.f. is $f(p)$ then the expectation E(P) is defined [Ref 8: p. 180] as follows :

$$E(P) = \int_{-\infty}^{\infty} p\, f(p)\, dp. \tag{A.5}$$

The mean of a random variable with the density function of Equation A.4 is then

$$E(P) = \int_0^{p_{max}} p \, \frac{2}{p_{max}^2} \, (p_{max} - p) \, dp = \frac{p_{max}}{3}. \qquad (A.6)$$

Now we will find the equation of the second variant of the triangular distribution with parameter p_{min} for the line through (1,a) and (p_{min} ,0) (see Figure 2). First, we will find the slope by using Equation A.1, that is

$$m = \frac{0 - a}{p_{min} - 1} = -\frac{a}{(p_{min} - 1)}.$$

where $0 \leq p_{min} \leq p < 1$ and $2 \leq a < \infty$. Using the point (1,a) and the slope $m = -\frac{a}{(p_{min} - 1)}$ in Equation A.2 then gives

$$y - a = -\frac{a}{(p_{min} - 1)} (p - 1).$$

Setting $y = f(p)$, we then have

$$f(p) = \frac{a}{p_{min} - 1} (p_{min} - p). \qquad (A.7)$$

Again we will find another equation for "a" by using the following property

$$\int_{p_{min}}^{1} \frac{a}{p_{min} - 1} (p_{min} - p) dp = 1,$$

which results in

$$a = \frac{2}{1 - p_{min}}.$$

Substituting "a" into Equation A.7, we have the density function

$$f(p) = -\frac{2(p_{min} - p)}{(p_{min} - 1)^2}, \qquad (A.8)$$

where $0 \leq p_{min} \leq p < 1$. The mean is

$$E(P) = \int_{p_{min}}^{1} p \, \frac{-2(p_{min} - p)}{(p_{min} - 1)^2} \, dp = \frac{1}{3} \, (p_{min} + 2). \qquad (A.9)$$

The above two variants of the triangular distribution are used as our priors to derive the posterior distribution based on Bayes' theorem in Chapter III of this thesis.

APPENDIX B. THE APL PROGRAM USED TO COMPUTE BAYESIAN INTERVALS WITH THE TRIANGULAR PRIOR DISTRIBUTION HAVING PARAMETER PMAX

```
    ∇ PMAXIMUM
   ⍝ THIS PROGRAM COMPUTES UPPER BOUNDS, LOWER BOUNDS AND BAYESIAN
   ⍝ CONFIDENCE INTERVALS TO DETERMINE THE NUMBER OF SAMPLES TO ESTIMATE
   ⍝ PROPORTIONS WITH 95 PERCENT CONFIDENCE USING PRIOR TRIANGULAR
   ⍝ DISTRIBUTION HAVING PARAMETER PMAX (0<P≤PMAX≤1).  IT ASKS THE
   ⍝ USER TO INPUT THE PARAMETER OF THE PRIOR TRIANGULAR DISTRIBUTION,
   ⍝ VECTOR OF SAMPLE SIZES, PROBABILITY VECTOR TO FIND BOUNDS
   ⍝ CONFIDENCE LEVEL.  ALSO IT USES THE PROGRAM BETA AS SUBROUTINE.

    ⎕←'ENTER SAMPLE SIZE'
    SS1←⎕
    ⎕←'ENTER BOUND OF PRIOR TRIANGULAR DISTRIBUTION'
    PM1←⎕
    SAY←0
   LOOP1:SAY←SAY+1
    ⎕←'ENTER PROBABILITY VECTOR FOR TESTING PLOW OR PUPPER'
    P←⎕
    ⎕←'ENTER 0.025 OR 0.975 RELATING TO THE CONFIDENCE LEVEL '
    INT←⎕
    K←0
   LOOP2:
    PLO←P[K+1]
    ALFHA1←(((SS1×PM1)+3)÷3)
    ALFHA2←(((SS1×PM1)+6)÷3)
    BETA1←(((3×SS1)+3)-(SS1×PM1))÷3
    N1←(((SS1×PM1)+3)÷((3×SS1)+6))
    PAR1←ALFHA1,BETA1
    PAR2←ALFHA2,BETA1
    TPAR←ALFHA1,ALFHA2,BETA1,N1
    B1←((PM1×(PAR1 BETA PLO))-(PM1×(PAR1 BETA 0)))
    B2←((N1×(PAR2 BETA PLO))-(N1×(PAR2 BETA 0)))
    B6←PM1×(PAR1 BETA PM1)
    B8←N1×(PAR2 BETA PM1)
    DE←B6-B8
    FLO←(B1-B2)÷DE
    K←K+1
    →(FLO≤INT)/LOOP2
    →(0.5<INT)/SONA
    'PLOW = ',⍕PLO
    PLO2←PLO
    →(1≤SAY)/LOOP1
   SONA:
    PLO1←PLO
    'PUPPER = ',⍕PLO1
    CI←PLO1-PLO2
    'SIZE OF BAYESIAN INTERVAL 2A = ',⍕CI
    ∇
```

APPENDIX C. THE APL PROGRAM USED TO COMPUTE BAYESIAN INTERVALS WITH THE TRIANGULAR PRIOR DISTRIBUTION HAVING PARAMETER PMIN

```
    ∇ PMINIMUM
  ⍝ THIS PROGRAM COMPUTES UPPER BOUNDS, LOWER BOUNDS AND BAYESIAN
  ⍝ CONFIDENCE INTERVALS TO DETERMINE THE NUMBER OF SAMPLES USING
  ⍝ PRIOR TRIANGULAR DISTRIBUTION (0≤PMIN≤P<1). IT ASK THE USER
  ⍝ TO INPUT THE VECTOR OF SAMPLE SIZE, THE PARAMETER OF PRIOR
  ⍝ TRIANGULAR DISTRIBUTION AND CONFIDENCE LEVEL. IT USES PROGRAM
  ⍝ BETA AS SUBROUTINE.

   ⎕←'ENTER SAMPLE SIZE'
   SS1←⎕
   ⎕←'ENTER BOUND OF PRIOR TRIANGULAR DISTRIBUTION'
   PM1←⎕
   SAY←0
  LOOP1:SAY←SAY+1
   ⎕←'ENTER PROBABILITY VECTOR FOR TESTING PLOW OR PUPPER'
   P←⎕
   ⎕←'ENTER 0.025 OR 0.975 RELATED TO THE CONFIDENCE LEVEL'
   INT←⎕
   K←0
  LOOP2:
   PLO←P[K+1]
   ALFHA3←(((SS1×PM1)+(2×SS1)+3)÷3)
   ALFHA4←(((SS1×PM1)+(2×SS1)+6)÷3)
   BETA2←((SS1+3-(SS1×PM1))÷3)
   N2←(((SS1×PM1)+(2×SS1)+3)÷((3×SS1)+6))
   PAR1←ALFHA3,BETA2
   PAR2←ALFHA4,BETA2
   TPAR←ALFHA3,ALFHA4,BETA2,N2
   B1←((PM1×(PAR1 BETA PLO))-(PM1×(PAR1 BETA PM1)))
   B2←((N2×(PAR2 BETA PLO))-(N2×(PAR2 BETA PM1)))
   B5←(PM1×(PAR1 BETA 1))
   B6←(PM1×(PAR1 BETA PM1))
   DE1←B5-B6
   B7←(N2×(PAR2 BETA 1))
   B8←(N2×(PAR2 BETA PM1))
   DE2←B7-B8
   DE←DE1-DE2
   PLO←(B1-B2)÷DE
   K←K+1
   →(PLO≤INT)/LOOP2
   →(0.5<INT)/SONA
   'PLOW = ',⍕PLO
   PLO2←PLO
   →(1≤SAY)/LOOP1
  SONA:
   PLO1←PLO
   'PUPPER = ',⍕PLO1
   CI←PLO1-PLO2
   'SIZE OF BAYESIAN INTERVAL 2A = ',⍕CI
    ∇
```

APPENDIX D. THE APL PROGRAM USED TO COMPUTE THE BETA DENSITY FUNCTION

```
    ∇ U←A BETA X;Y;W;N;OD;EV;Z;I
   ⍝ 12/27/86 EVALUATES THE BETA CDF, PARAMETERS A, AT VECTOR X USING THE
   ⍝ BOUVER-BARGMAN CONTINUED FRACTION AT DEPTH VARYING FROM 7 TO 21.
   ⍝ 11TH ANNUAL SYMPOSIUM ON THE INTERFACE OF COMPUTER SCIENCE AND
   ⍝ STATISTICS, 1978, P 325. BECAUSE OF THE RANGE OF !, +/A≤255. SEEMS TO
   ⍝ GIVE A GOOD 8 OR MORE DECIMALS.
    Y←X≤(A[1]÷+/A)
    U←(ρ,X)ρ0
    N←7++/(⌈/A)>(2×⍳4),10×⍳10
    →((+/Y)=0)/FLIP
    W←Y/X←,X
    OD←W∘.×((⍳N)×A[2]-⍳N)÷×/(N,2)ρA[1]+⍳2×I←N
    EV←-W∘.×(×⌿((2,N)ρ(A[1]+0,⍳N-1),(+/A)+0,⍳N-1))÷×/(N,2)ρA[1]+0,⍳(2×N-Z←1)
  L:Z←1+EV[;I]÷1+OD[;I]÷Z
    →((I←I-1)>0)/L
    U[Y/⍳ρU]←(÷Z)×(A[1]!¯1++/A)×(W*A[1])×(1-W)*A[2]
    →((+/Y)=ρX)/0
  FLIP:A←⌽A
    W←1-(~Y)/X
    OD←W∘.×((⍳N)×A[2]-⍳N)÷×/(N,2)ρA[1]+⍳2×I←N
    EV←-W∘.×(×⌿((2,N)ρ(A[1]+0,⍳N-1),(+/A)+0,⍳N-1))÷×/(N,2)ρA[1]+0,⍳(2×N-Z←1)
  L1:Z←1+EV[;I]÷1+OD[;I]÷Z
    →((I←I-1)>0)/L1
    U[(~Y)/⍳ρU]←1-(÷Z)×(A[1]!¯1++/A)×(W*A[1])×(1-W)*A[2]
    ∇
```

APPENDIX E. TABLES THAT CAN BE USED TO DETERMINE SAMPLE SIZES BY USING THE PRIOR TRIANGULAR DISTRIBUTION WITH VARIOUS PMIN PARAMETERS

Table 5. SAMPLE SIZES AND BAYESIAN INTERVALS USING THE TRIANGULAR PRIOR DISTRIBUTION WITH PARAMETER PMIN=0.0

Sample Size n	Lower Bound	Upper Bound	Bayesian Interval Size = 2A
1	0.2110	0.9741	0.7631
2	0.2516	0.9603	0.7087
3	0.2836	0.9473	0.6637
4	0.3095	0.9353	0.6258
5	0.3310	0.9245	0.5935
6	0.3491	0.9148	0.5656
7	0.3647	0.9060	0.5412
8	0.3783	0.8980	0.5197
9	0.3903	0.8907	0.5005
10	0.4009	0.8841	0.4832
20	0.4666	0.8393	0.3728
30	0.4999	0.8143	0.3144
40	0.5209	0.7978	0.2769
50	0.5356	0.7858	0.2502
60	0.5466	0.7767	0.2300
70	0.5553	0.7694	0.2140
80	0.5624	0.7634	0.2010
90	0.5682	0.7583	0.1901
100	0.5732	0.7540	0.1807
110	0.5775	0.7502	0.1727
120	0.5813	0.7469	0.1656
130	0.5846	0.7440	0.1593
140	0.5876	0.7413	0.1537
150	0.5903	0.7389	0.1487
160	0.5927	0.7368	0.1441

Table 6. **SAMPLE SIZES AND BAYESIAN INTERVALS USING THE TRIANGULAR PRIOR DISTRIBUTION WITH PARAMETER PMIN=0.1**

Sample Size n	Lower Bound	Upper Bound	Bayesian Interval Size = 2A
1	0.2738	0.9779	0.7041
2	0.3029	0.9667	0.6638
3	0.3287	0.9560	0.6273
4	0.3511	0.9460	0.5949
5	0.3705	0.9370	0.5665
6	0.3873	0.9287	0.5414
7	0.4021	0.9212	0.5191
8	0.4151	0.9143	0.4992
9	0.4266	0.9080	0.4814
10	0.4370	0.9022	0.4652
20	0.5019	0.8625	0.3606
30	0.5351	0.8397	0.3046
40	0.5558	0.8246	0.2688
50	0.5705	0.8135	0.2430
60	0.5816	0.8050	0.2234
70	0.5902	0.7981	0.2079
80	0.5972	0.7925	0.1952
90	0.6031	0.7877	0.1846
100	0.6081	0.7837	0.1756
110	0.6123	0.7801	0.1678
120	0.6161	0.7770	0.1609
130	0.6194	0.7742	0.1548
140	0.6223	0.7717	0.1494
150	0.6250	0.7694	0.1445
160	0.6274	0.7574	0.1400
170	0.6296	0.7655	0.1359
180	0.6316	0.7638	0.1322
190	0.6334	0.7622	0.1288
200	0.6351	0.7607	0.1256
210	0.6367	0.7593	0.1226
220	0.6382	0.7580	0.1199
230	0.6395	0.7568	0.1173
240	0.6408	0.7557	0.1148
250	0.6420	0.7546	0.1126
260	0.6432	0.7536	0.1104
270	0.6442	0.7526	0.1084
280	0.6453	0.7517	0.1065

Table 7. SAMPLE SIZES AND BAYESIAN INTERVALS USING THE TRIANGULAR PRIOR DISTRIBUTION WITH PARAMETER PMIN = 0.2

Sample Size n	Lower Bound	Upper Bound	Bayesian Interval Size = 2A
1	0.3459	0.9815	0.6356
2	0.3653	0.9726	0.6073
3	0.3839	0.9640	0.5801
4	0.4012	0.9559	0.5547
5	0.4171	0.9485	0.5314
6	0.4314	0.9416	0.5102
7	0.4444	0.9353	0.4909
8	0.4561	0.9295	0.4734
9	0.4667	0.9242	0.4575
10	0.4764	0.9192	0.4428
20	0.5388	0.8845	0.3457
30	0.5715	0.8642	0.2927
40	0.5921	0.8504	0.2583
50	0.6066	0.8403	0.2337
60	0.6175	0.8325	0.2150
70	0.6260	0.8262	0.2002
80	0.6330	0.8210	0.1880
90	0.6385	0.8166	0.1781
100	0.6435	0.8128	0.1694
110	0.6477	0.8095	0.1618
120	0.6514	0.8066	0.1552
130	0.6546	0.8040	0.1493
140	0.6575	0.8016	0.1441
150	0.6601	0.7995	0.1393
160	0.6625	0.7976	0.1350
170	0.6647	0.7958	0.1311
180	0.6666	0.7942	0.1275
190	0.6684	0.7926	0.1242
200	0.6701	0.7912	0.1211

Table 8. SAMPLE SIZES AND BAYESIAN INTERVALS USING THE TRIANGULAR PRIOR DISTRIBUTION WITH PARAMETER PMIN=0.3

Sample Size n	Lower Bound	Upper Bound	Bayesian Interval Size = 2A
1	0.4227	0.9848	0.5621
2	0.4352	0.9780	0.5428
3	0.4477	0.9713	0.5236
4	0.4599	0.9649	0.5050
5	0.4717	0.9590	0.4873
6	0.4828	0.9535	0.4707
7	0.4932	0.9483	0.4551
8	0.5029	0.9436	0.4407
9	0.5119	0.9392	0.4273
10	0.5203	0.9351	0.4148
20	0.5778	0.9055	0.3277
30	0.6093	0.8877	0.2784
40	0.6294	0.8754	0.2460
50	0.6436	0.8664	0.2228
60	0.6542	0.8593	0.2051
70	0.6625	0.8536	0.1911
80	0.6693	0.8488	0.1795
90	0.6747	0.8448	0.1702
100	0.6795	0.8413	0.1619
110	0.6836	0.8383	0.1547
120	0.6872	0.8356	0.1483
130	0.6904	0.8332	0.1427
140	0.6933	0.8310	0.1377
150	0.6958	0.8290	0.1332
160	0.6981	0.8272	0.1291
170	0.7002	0.8256	0.1254
180	0.7022	0.8241	0.1219
190	0.7039	0.8227	0.1188
200	0.7056	0.8214	0.1158
210	0.7071	0.8202	0.1131
220	0.7085	0.8190	0.1106
230	0.7098	0.8180	0.1082
240	0.7110	0.8170	0.1059

Table 9. **SAMPLE SIZES AND BAYESIAN INTERVALS USING THE TRIANGULAR PRIOR DISTRIBUTION WITH PARAMETER PMIN=0.4**

Sample Size n	Lower Bound	Upper Bound	Bayesian Interval Size = 2A
1	0.5022	0.9878	0.4856
2	0.5099	0.9828	0.4729
3	0.5179	0.9778	0.4599
4	0.5259	0.9731	0.4472
5	0.5339	0.9685	0.4346
6	0.5417	0.9643	0.4226
7	0.5492	0.9602	0.4110
8	0.5565	0.9565	0.4000
9	0.5634	0.9530	0.3896
10	0.5701	0.9497	0.3796
20	0.6196	0.9253	0.3057
30	0.6489	0.9101	0.2612
40	0.6681	0.8995	0.2314
50	0.6817	0.8915	0.2098
60	0.6919	0.8852	0.1933
70	0.7000	0.8801	0.1801
80	0.7065	0.8759	0.1694
90	0.7116	0.8723	0.1608
100	0.7163	0.8692	0.1529
110	0.7203	0.8664	0.1461
120	0.7238	0.8640	0.1402
130	0.7269	0.8618	0.1349
140	0.7296	0.8598	0.1302
150	0.7321	0.8580	0.1259
160	0.7343	0.8564	0.1220
170	0.7364	0.8549	0.1185
180	0.7382	0.8535	0.1152
190	0.7399	0.8522	0.1123
200	0.7415	0.8510	0.1095
210	0.7430	0.8499	0.1069
220	0.7443	0.8488	0.1045
230	0.7456	0.8479	0.1023
240	0.7468	0.8469	0.1002
250	0.7479	0.8461	0.0982
260	0.7490	0.8453	0.0963
270	0.7499	0.8445	0.0945

Table 10. SAMPLE SIZES AND BAYESIAN INTERVALS USING THE TRIANGULAR PRIOR DISTRIBUTION WITH PARAMETER PMIN = 0.5

Sample Size n	Lower Bound	Upper Bound	Bayesian Interval Size = 2A
1	0.5833	0.9905	0.4072
2	0.5879	0.9871	0.3992
3	0.5927	0.9836	0.3909
4	0.5975	0.9802	0.3827
5	0.6025	0.9770	0.3745
6	0.6074	0.9739	0.3665
7	0.6124	0.9709	0.3585
8	0.6172	0.9681	0.3509
9	0.6220	0.9654	0.3434
10	0.6267	0.9629	0.3362
20	0.6654	0.9437	0.2783
30	0.6911	0.9312	0.2401
40	0.7086	0.9223	0.2137
50	0.7213	0.9155	0.1942
60	0.7310	0.9101	0.1791
70	0.7386	0.9057	0.1671
80	0.7448	0.9020	0.1572
90	0.7500	0.8989	0.1489
100	0.7544	0.8960	0.1416
110	0.7582	0.8937	0.1355
120	0.7611	0.8916	0.1305
130	0.7641	0.8897	0.1256
140	0.7667	0.8879	0.1212
150	0.7691	0.8863	0.1172
160	0.7712	0.8848	0.1136
170	0.7732	0.8835	0.1103
180	0.7750	0.8823	0.1073
190	0.7766	0.8811	0.1045
200	0.7781	0.8800	0.1019
210	0.7795	0.8790	0.0996
220	0.7808	0.8781	0.0973
230	0.7820	0.8772	0.0952
240	0.7831	0.8764	0.0933
250	0.7842	0.8756	0.0914
260	0.7852	0.8749	0.0897
270	0.7861	0.8742	0.0880
280	0.7870	0.8735	0.0865
290	0.7879	0.8729	0.0850
300	0.7887	0.8723	0.0836

Table 11. SAMPLE SIZES AND BAYESIAN INTERVALS USING THE TRIANGULAR PRIOR DISTRIBUTION WITH PARAMETER PMIN = 0.6

Sample Size n	Lower Bound	Upper Bound	Bayesian Interval Size = 2A
1	0.6656	0.9930	0.3274
2	0.6681	0.9908	0.3227
3	0.6707	0.9886	0.3179
4	0.6734	0.9864	0.3130
5	0.6762	0.9842	0.3080
6	0.6790	0.9821	0.3031
7	0.6818	0.9801	0.2983
8	0.6847	0.9782	0.2935
9	0.6876	0.9763	0.2887
10	0.6904	0.9746	0.2842
20	0.7167	0.9604	0.2437
30	0.7369	0.9508	0.2139
40	0.7518	0.9437	0.1919
50	0.7631	0.9382	0.1751
60	0.7718	0.9338	0.1620
70	0.7788	0.9301	0.1513
80	0.7845	0.9271	0.1426
90	0.7893	0.9244	0.1351
100	0.7934	0.9221	0.1287
110	0.7969	0.9200	0.1231
120	0.8000	0.9182	0.1182
130	0.8028	0.9165	0.1137
140	0.8052	0.9150	0.1098
150	0.8074	0.9137	0.1063
160	0.8093	0.9124	0.1031
170	0.8108	0.9113	0.1005
180	0.8125	0.9103	0.0978
190	0.8140	0.9093	0.0953
200	0.8154	0.9083	0.0929
210	0.8167	0.9075	0.0907
220	0.8180	0.9067	0.0887
230	0.8191	0.9059	0.0868
240	0.8201	0.9052	0.0850
250	0.8211	0.9045	0.0833
260	0.8221	0.9038	0.0818
270	0.8230	0.9032	0.0803
280	0.8238	0.9026	0.0788
290	0.8246	0.9021	0.0775
300	0.8253	0.9015	0.0762
350	0.8286	0.8992	0.0707
400	0.8311	0.8973	0.0662

Table 12. SAMPLE SIZES AND BAYESIAN INTERVALS USING THE TRIANGULAR PRIOR DISTRIBUTION WITH PARAMETER PMIN=0.7

Sample Size n	Lower Bound	Upper Bound	Bayesian Interval Size = 2A
1	0.7486	0.9951	0.2465
2	0.7498	0.9939	0.2441
3	0.7510	0.9927	0.2417
4	0.7523	0.9915	0.2392
5	0.7537	0.9902	0.2365
6	0.7551	0.9890	0.2339
7	0.7565	0.9878	0.2313
8	0.7579	0.9867	0.2288
9	0.7593	0.9855	0.2262
10	0.7608	0.9844	0.2236
20	0.7754	0.9752	0.1998
30	0.7885	0.9684	0.1799
40	0.7994	0.9633	0.1639
50	0.8082	0.9592	0.1510
60	0.8155	0.9559	0.1404
70	0.8214	0.9531	0.1317
80	0.8263	0.9507	0.1244
90	0.8306	0.9486	0.1180
100	0.8342	0.9468	0.1126
110	0.8373	0.9451	0.1078
120	0.8401	0.9436	0.1035
130	0.8425	0.9424	0.0999
140	0.8447	0.9411	0.0964
150	0.8466	0.9400	0.0934
160	0.8479	0.9391	0.0912
170	0.8496	0.9381	0.0885
180	0.8511	0.9373	0.0861
190	0.8525	0.9364	0.0839
200	0.8538	0.9357	0.0819
210	0.8550	0.9350	0.0800
220	0.8561	0.9343	0.0782
230	0.8571	0.9336	0.0765
240	0.8581	0.9330	0.0749
250	0.8590	0.9325	0.0735
260	0.8599	0.9319	0.0721
270	0.8607	0.9314	0.0708
280	0.8614	0.9309	0.0695
290	0.8621	0.9305	0.0683
300	0.8628	0.9300	0.0672
350	0.8658	0.9281	0.0623
400	0.8681	0.9265	0.0583
450	0.8701	0.9251	0.0551
500	0.8717	0.9240	0.0523

Table 13. SAMPLE SIZES AND BAYESIAN INTERVALS USING THE TRIANGULAR PRIOR DISTRIBUTION WITH PARAMETER PMIN = 0.8

Sample Size n	Lower Bound	Upper Bound	Bayesian Interval Size = 2A
1	0.8321	0.9970	0.1649
2	0.8325	0.9965	0.1640
3	0.8330	0.9960	0.1630
4	0.8335	0.9954	0.1619
5	0.8340	0.9949	0.1609
6	0.8345	0.9943	0.1598
7	0.8351	0.9938	0.1587
8	0.8356	0.9932	0.1576
9	0.8361	0.9927	0.1566
10	0.8367	0.9922	0.1555
20	0.8426	0.9874	0.1448
30	0.8486	0.9836	0.1350
40	0.8543	0.9804	0.1261
50	0.8596	0.9779	0.1183
60	0.8642	0.9757	0.1115
70	0.8683	0.9738	0.1055
80	0.8719	0.9723	0.1004
90	0.8751	0.9709	0.0958
100	0.8779	0.9696	0.0917
110	0.8804	0.9685	0.0881
120	0.8826	0.9674	0.0848
130	0.8846	0.9665	0.0819
140	0.8853	0.9657	0.0803
150	0.8871	0.9649	0.0778
160	0.8887	0.9641	0.0754
170	0.8901	0.9634	0.0733
180	0.8914	0.9628	0.0713
190	0.8926	0.9622	0.0695
200	0.8938	0.9616	0.0678
210	0.8948	0.9611	0.0663
220	0.8957	0.9606	0.0648
230	0.8966	0.9601	0.0634
240	0.8975	0.9596	0.0621
250	0.8983	0.9592	0.0609
260	0.8990	0.9588	0.0598
270	0.8997	0.9584	0.0587
280	0.9004	0.9580	0.0577
290	0.9010	0.9577	0.0567
300	0.9016	0.9573	0.0558
350	0.9041	0.9559	0.0517
400	0.9062	0.9546	0.0484
450	0.9078	0.9536	0.0457
500	0.9093	0.9527	0.0434
550	0.9105	0.9519	0.0414
600	0.9115	0.9512	0.0397
650	0.9124	0.9506	0.0381
700	0.9132	0.9500	0.0368

APPENDIX F. TABLES THAT CAN BE USED TO DETERMINE SAMPLE SIZES BY USING THE PRIOR TRIANGULAR DISTRIBUTION WITH VARIOUS PMAX PARAMETERS

Table 14. SAMPLE SIZES AND BAYESIAN INTERVALS USING THE TRIANGULAR PRIOR DISTRIBUTION WITH PARAMETER PMAX=1.0

Sample Size n	Lower Bound	Upper Bound	Bayesian Interval Size = 2A
1	0.0259	0.7890	0.7631
2	0.0397	0.7484	0.7087
3	0.0527	0.7164	0.6637
4	0.0647	0.6905	0.6258
5	0.0755	0.6690	0.5935
6	0.0852	0.6509	0.5656
7	0.0940	0.6353	0.5412
8	0.1020	0.6217	0.5197
9	0.1093	0.6097	0.5005
10	0.1159	0.5991	0.4832
20	0.1607	0.5334	0.3728
30	0.1857	0.5001	0.3144
40	0.2022	0.4791	0.2769
50	0.2142	0.4644	0.2502
60	0.2233	0.4534	0.2300
70	0.2306	0.4447	0.2140
80	0.2366	0.4376	0.2010
90	0.2417	0.4318	0.1901
100	0.2460	0.4268	0.1807
110	0.2498	0.4225	0.1727
120	0.2531	0.4187	0.1656
130	0.2560	0.4154	0.1593
140	0.2587	0.4124	0.1537
150	0.2611	0.4097	0.1487
160	0.2632	0.4073	0.1441

Table 15. SAMPLE SIZES AND BAYESIAN INTERVALS USING THE TRIANGULAR PRIOR DISTRIBUTION WITH PARAMETER PMAX=0.9

Sample Size n	Lower Bound	Upper Bound	Bayesian Interval Size = 2A
1	0.0221	0.7262	0.7041
2	0.0339	0.6971	0.6638
3	0.0440	0.6713	0.6273
4	0.0540	0.6489	0.5949
5	0.0630	0.6295	0.5665
6	0.0713	0.6127	0.5414
7	0.0788	0.5979	0.5191
8	0.0857	0.5849	0.4992
9	0.0920	0.5734	0.4814
10	0.0978	0.5630	0.4652
20	0.1375	0.4981	0.3606
30	0.1603	0.4649	0.3046
40	0.1754	0.4442	0.2688
50	0.1865	0.4295	0.2430
60	0.1950	0.4184	0.2234
70	0.2019	0.4098	0.2079
80	0.2075	0.4028	0.1952
90	0.2123	0.3969	0.1846
100	0.2163	0.3919	0.1756
110	0.2199	0.3877	0.1678
120	0.2230	0.3839	0.1609
130	0.2258	0.3806	0.1548
140	0.2283	0.3777	0.1494
150	0.2306	0.3750	0.1445
160	0.2326	0.3726	0.1400
170	0.2345	0.3704	0.1359
180	0.2362	0.3684	0.1322
190	0.2378	0.3666	0.1288
200	0.2393	0.3649	0.1256
210	0.2407	0.3633	0.1226
220	0.2420	0.3618	0.1199
230	0.2432	0.3605	0.1173
240	0.2443	0.3592	0.1148
250	0.2454	0.3580	0.1126
260	0.2464	0.3568	0.1104
270	0.2474	0.3558	0.1084
280	0.2483	0.3547	0.1065

Table 16. SAMPLE SIZES AND BAYESIAN INTERVALS USING THE TRIANGULAR PRIOR DISTRIBUTION WITH PARAMETER PMAX=0.8

Sample Size n	Lower Bound	Upper Bound	Bayesian Interval Size = 2A
1	0.0185	0.6541	0.6356
2	0.0274	0.6347	0.6073
3	0.0360	0.6161	0.5801
4	0.0441	0.5988	0.5547
5	0.0515	0.5829	0.5314
6	0.0584	0.5686	0.5102
7	0.0647	0.5556	0.4909
8	0.0705	0.5439	0.4734
9	0.0758	0.5333	0.4575
10	0.0808	0.5236	0.4428
20	0.1155	0.4612	0.3457
30	0.1358	0.4285	0.2927
40	0.1496	0.4079	0.2583
50	0.1597	0.3934	0.2337
60	0.1675	0.3825	0.2150
70	0.1738	0.3740	0.2002
80	0.1790	0.3670	0.1880
90	0.1834	0.3615	0.1781
100	0.1872	0.3565	0.1694
110	0.1905	0.3523	0.1618
120	0.1934	0.3486	0.1552
130	0.1960	0.3454	0.1493
140	0.1984	0.3425	0.1441
150	0.2005	0.3399	0.1393
160	0.2024	0.3375	0.1350
170	0.2042	0.3353	0.1311
180	0.2058	0.3334	0.1275
190	0.2074	0.3316	0.1242
200	0.2088	0.3299	0.1211

Table 17. SAMPLE SIZES AND BAYESIAN INTERVALS USING THE TRIANGULAR PRIOR DISTRIBUTION WITH PARAMETER PMAX = 0.7

Sample Size n	Lower Bound	Upper Bound	Bayesian Interval Size = 2A
1	0.0152	0.5773	0.5621
2	0.0220	0.5648	0.5428
3	0.0287	0.5523	0.5236
4	0.0351	0.5401	0.5050
5	0.0410	0.5283	0.4873
6	0.0465	0.5172	0.4707
7	0.0517	0.5068	0.4551
8	0.0564	0.4971	0.4407
9	0.0608	0.4881	0.4273
10	0.0649	0.4797	0.4148
20	0.0945	0.4222	0.3277
30	0.1123	0.3907	0.2784
40	0.1246	0.3706	0.2460
50	0.1336	0.3564	0.2228
60	0.1407	0.3458	0.2051
70	0.1464	0.3375	0.1911
80	0.1512	0.3307	0.1795
90	0.1552	0.3254	0.1702
100	0.1587	0.3205	0.1619
110	0.1617	0.3164	0.1547
120	0.1644	0.3128	0.1483
130	0.1668	0.3096	0.1427
140	0.1690	0.3067	0.1377
150	0.1710	0.3042	0.1332
160	0.1728	0.3019	0.1291
170	0.1744	0.2998	0.1254
180	0.1759	0.2978	0.1219
190	0.1773	0.2961	0.1188
200	0.1786	0.2944	0.1158
210	0.1798	0.2929	0.1131
220	0.1810	0.2915	0.1106
230	0.1820	0.2902	0.1082
240	0.1830	0.2890	0.1059

Table 18. SAMPLE SIZES AND BAYESIAN INTERVALS USING THE TRIANGULAR PRIOR DISTRIBUTION WITH PARAMETER PMAX=0.6

Sample Size n	Lower Bound	Upper Bound	Bayesian Interval Size = 2A
1	0.0122	0.4978	0.4856
2	0.0172	0.4901	0.4729
3	0.0222	0.4821	0.4599
4	0.0269	0.4741	0.4472
5	0.0315	0.4661	0.4346
6	0.0357	0.4583	0.4226
7	0.0398	0.4508	0.4110
8	0.0435	0.4435	0.4000
9	0.0470	0.4366	0.3896
10	0.0503	0.4299	0.3796
20	0.0747	0.3804	0.3057
30	0.0899	0.3511	0.2612
40	0.1005	0.3319	0.2314
50	0.1085	0.3183	0.2098
60	0.1148	0.3081	0.1933
70	0.1199	0.3000	0.1801
80	0.1241	0.2935	0.1694
90	0.1277	0.2884	0.1608
100	0.1308	0.2837	0.1529
110	0.1336	0.2797	0.1461
120	0.1360	0.2762	0.1402
130	0.1382	0.2731	0.1349
140	0.1402	0.2704	0.1302
150	0.1420	0.2679	0.1259
160	0.1436	0.2657	0.1220
170	0.1451	0.2636	0.1185
180	0.1465	0.2618	0.1152
190	0.1478	0.2601	0.1123
200	0.1490	0.2585	0.1095
210	0.1501	0.2570	0.1069
220	0.1512	0.2557	0.1045
230	0.1521	0.2544	0.1023
240	0.1531	0.2532	0.1002
250	0.1539	0.2521	0.0982
260	0.1547	0.2510	0.0963
270	0.1555	0.2501	0.0945

Table 19. SAMPLE SIZES AND BAYESIAN INTERVALS USING THE TRIANGULAR PRIOR DISTRIBUTION WITH PARAMETER PMAX = 0.5

Sample Size n	Lower Bound	Upper Bound	Bayesian Interval Size = 2A
1	0.0095	0.4167	0.4072
2	0.0129	0.4121	0.3992
3	0.0164	0.4073	0.3909
4	0.0198	0.4025	0.3827
5	0.0230	0.3975	0.3745
6	0.0261	0.3926	0.3665
7	0.0291	0.3876	0.3585
8	0.0319	0.3828	0.3509
9	0.0346	0.3780	0.3434
10	0.0371	0.3733	0.3362
20	0.0563	0.3346	0.2783
30	0.0688	0.3089	0.2401
40	0.0777	0.2914	0.2137
50	0.0845	0.2787	0.1942
60	0.0899	0.2690	0.1791
70	0.0943	0.2614	0.1671
80	0.0980	0.2552	0.1572
90	0.1011	0.2500	0.1489
100	0.1040	0.2456	0.1416
110	0.1062	0.2422	0.1360
120	0.1084	0.2389	0.1305
130	0.1103	0.2359	0.1256
140	0.1121	0.2333	0.1212
150	0.1137	0.2309	0.1172
160	0.1152	0.2288	0.1136
170	0.1165	0.2268	0.1103
180	0.1177	0.2250	0.1073
190	0.1189	0.2234	0.1045
200	0.1200	0.2219	0.1019
210	0.1210	0.2205	0.0996
220	0.1219	0.2192	0.0973
230	0.1228	0.2180	0.0952
240	0.1236	0.2169	0.0933
250	0.1244	0.2158	0.0914
260	0.1251	0.2148	0.0897
270	0.1258	0.2139	0.0880
280	0.1265	0.2130	0.0865
290	0.1271	0.2121	0.0850
300	0.1277	0.2113	0.0836

Table 20. SAMPLE SIZES AND BAYESIAN INTERVALS USING THE TRIANGULAR PRIOR DISTRIBUTION WITH PARAMETER PMAX = 0.4

Sample Size n	Lower Bound	Upper Bound	Bayesian Interval Size = 2A
1	0.0070	0.3344	0.3274
2	0.0082	0.3309	0.3227
3	0.0114	0.3293	0.3179
4	0.0136	0.3266	0.3130
5	0.0158	0.3238	0.3080
6	0.0179	0.3210	0.3031
7	0.0199	0.3182	0.2983
8	0.0218	0.3153	0.2935
9	0.0237	0.3124	0.2887
10	0.0254	0.3096	0.2842
20	0.0396	0.2833	0.2437
30	0.0492	0.2631	0.2139
40	0.0563	0.2482	0.1919
50	0.0618	0.2369	0.1751
60	0.0662	0.2282	0.1620
70	0.0699	0.2212	0.1513
80	0.0729	0.2155	0.1426
90	0.0756	0.2107	0.1351
100	0.0779	0.2066	0.1287
110	0.0800	0.2031	0.1231
120	0.0818	0.2000	0.1182
130	0.0835	0.1972	0.1137
140	0.0850	0.1948	0.1098
150	0.0863	0.1926	0.1063
160	0.0875	0.1910	0.1035
170	0.0887	0.1892	0.1005
180	0.0897	0.1875	0.0978
190	0.0907	0.1860	0.0953
200	0.0917	0.1846	0.0929
210	0.0925	0.1833	0.0907
220	0.0933	0.1820	0.0887
230	0.0941	0.1809	0.0868
240	0.0948	0.1799	0.0850
250	0.0955	0.1789	0.0833
260	0.0962	0.1779	0.0818
270	0.0968	0.1770	0.0803
280	0.0974	0.1762	0.0788
290	0.0979	0.1754	0.0775
300	0.0985	0.1747	0.0762
350	0.1008	0.1714	0.0707
400	0.1027	0.1689	0.0662

Table 21. **SAMPLE SIZES AND BAYESIAN INTERVALS USING THE TRIANGULAR PRIOR DISTRIBUTION WITH PARAMETER PMAX = 0.3**

Sample Size n	Lower Bound	Upper Bound	Bayesian Interval Size = 2A
1	0.0049	0.2449	0.2465
2	0.0061	0.2502	0.2441
3	0.0073	0.2490	0.2417
4	0.0085	0.2477	0.2392
5	0.0098	0.2463	0.2365
6	0.0110	0.2449	0.2339
7	0.0122	0.2435	0.2313
8	0.0133	0.2421	0.2288
9	0.0145	0.2407	0.2262
10	0.0136	0.2392	0.2236
20	0.0248	0.2246	0.1998
30	0.0316	0.2115	0.1199
40	0.0367	0.2006	0.1639
50	0.0408	0.1918	0.1510
60	0.0441	0.1845	0.1404
70	0.0469	0.1786	0.1317
80	0.0493	0.1737	0.1244
90	0.0514	0.1694	0.1180
100	0.0632	0.1658	0.1126
110	0.0549	0.1627	0.1078
120	0.0564	0.1599	0.1035
130	0.0576	0.1575	0.0999
140	0.0589	0.1553	0.0964
150	0.0600	0.1534	0.0934
160	0.0609	0.1521	0.0912
170	0.0619	0.1504	0.0885
180	0.0627	0.1489	0.0861
190	0.0636	0.1475	0.0839
200	0.0643	0.1462	0.0819
210	0.0650	0.1450	0.0800
220	0.0657	0.1439	0.0782
230	0.0664	0.1429	0.0765
240	0.0670	0.1419	0.0749
250	0.0675	0.1410	0.0735
260	0.0681	0.1401	0.0721
270	0.0686	0.1393	0.0708
280	0.0691	0.1386	0.0695
290	0.0695	0.1379	0.0683
300	0.0700	0.1372	0.0672
350	0.0719	0.1342	0.0623
400	0.0735	0.1319	0.0583
450	0.0749	0.1299	0.0551
500	0.0760	0.1283	0.0523

Table 22. SAMPLE SIZES AND BAYESIAN INTERVALS USING THE TRIANGULAR PRIOR DISTRIBUTION WITH PARAMETER PMAX = 0.2

Sample Size n	Lower Bound	Upper Bound	Bayesian Interval Size = 2A
1	0.0030	0.1679	0.1649
2	0.0035	0.1675	0.1640
3	0.0040	0.1670	0.1630
4	0.0046	0.1665	0.1619
5	0.0051	0.1660	0.1609
6	0.0057	0.1655	0.1598
7	0.0062	0.1649	0.1587
8	0.0068	0.1644	0.1576
9	0.0073	0.1639	0.1566
10	0.0078	0.1633	0.1555
20	0.0126	0.1574	0.1448
30	0.0164	0.1514	0.1350
40	0.0196	0.1457	0.1261
50	0.0221	0.1404	0.1183
60	0.0243	0.1358	0.1115
70	0.0262	0.1317	0.1055
80	0.0277	0.1281	0.1004
90	0.0291	0.1249	0.0958
100	0.0304	0.1221	0.0917
110	0.0315	0.1196	0.0881
120	0.0326	0.1174	0.0848
130	0.0334	0.1166	0.0832
140	0.0343	0.1147	0.0803
150	0.0351	0.1129	0.0778
160	0.0359	0.1113	0.0754
170	0.0366	0.1099	0.0733
180	0.0372	0.1086	0.0713
190	0.0378	0.1074	0.0695
200	0.0384	0.1062	0.0678
210	0.0389	0.1052	0.0663
220	0.0394	0.1043	0.0648
230	0.0399	0.1034	0.0634
240	0.0404	0.1025	0.0621
250	0.0408	0.1017	0.0609
260	0.0412	0.1010	0.0598
270	0.0416	0.1003	0.0587
280	0.0420	0.0996	0.0577
290	0.0423	0.0990	0.0567
300	0.0427	0.0984	0.0558
350	0.0441	0.0959	0.0517
400	0.0454	0.0938	0.0484
450	0.0464	0.0922	0.0457
500	0.0473	0.0907	0.0434
550	0.0481	0.0895	0.0414
600	0.0488	0.0885	0.0397
650	0.0494	0.0876	0.0381
700	0.0500	0.0868	0.0368

APPENDIX G. GRAPHS THAT CAN BE USED TO DETERMINE SAMPLE SIZES BY USING PRIOR TRIANGULAR DISTRIBUTION WITH VARIOUS PARAMETERS PMIN OR PMAX

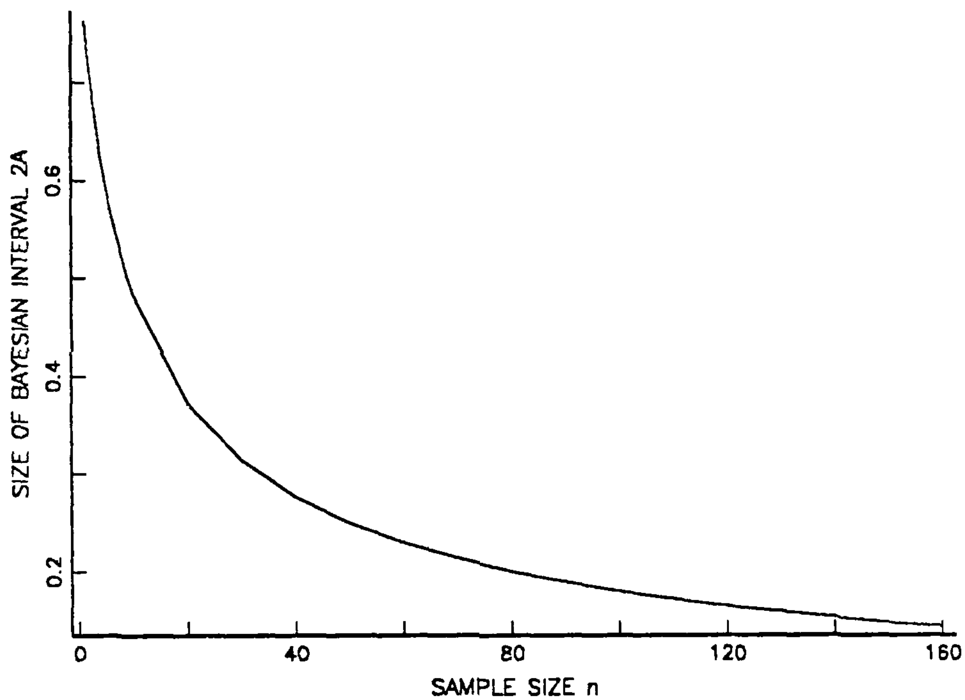

Figure 9. Number of Samples vs the Size of the 95% Bayesian Interval with a Triangular Prior Distribution with Pmin = 0 or Pmax = 1

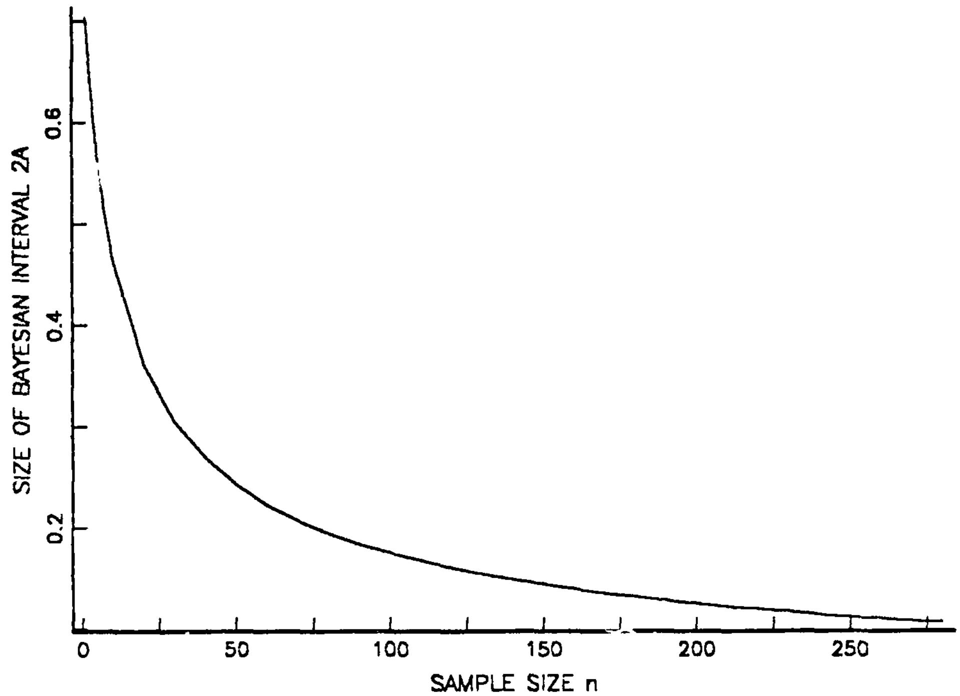

Figure 10. Number of Samples vs The Size of the 95% Bayesian Interval with a Triangular Prior Distribution with Pmin = 0.1 or Pmax = 0.9

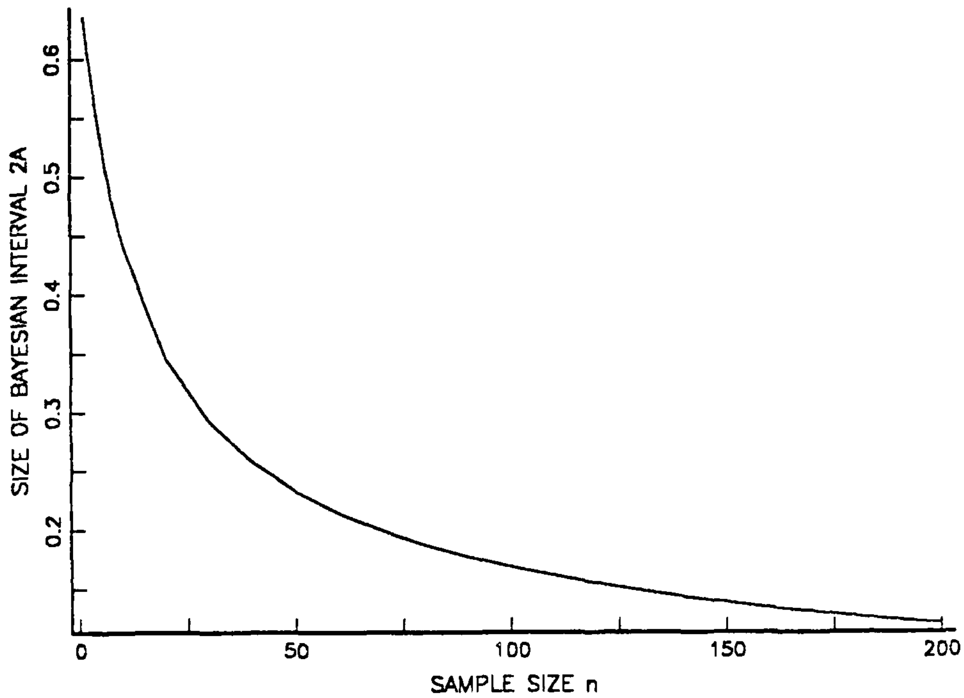

Figure 11. Number of Samples vs The Size of the 95% Bayesian Interval with a Triangular Prior Distribution with Pmin = 0.2 or Pmax = 0.8

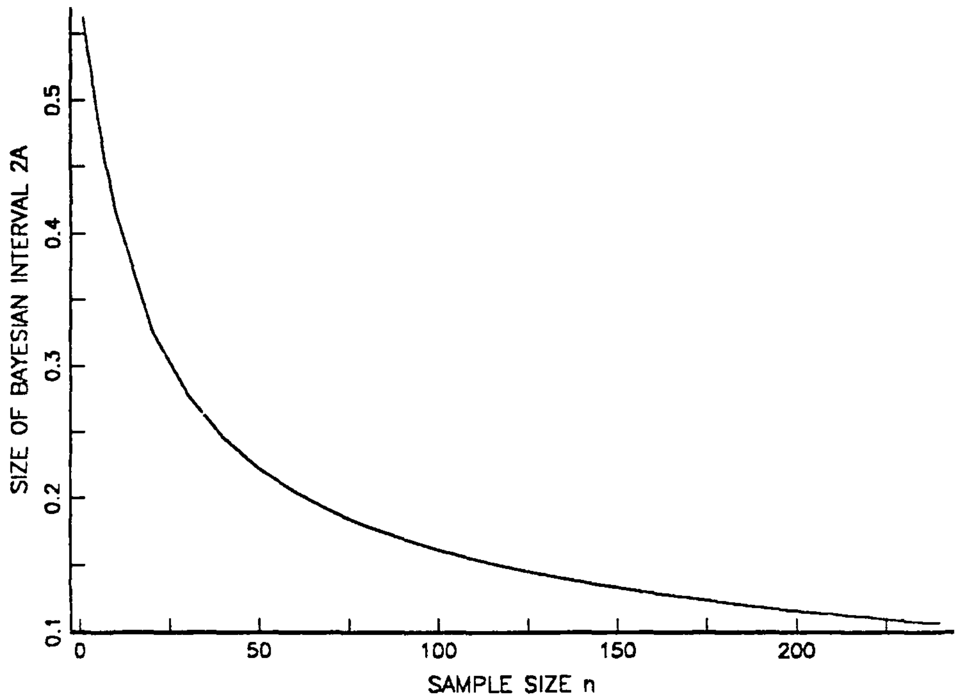

Figure 12. Number of Samples vs The Size of the 95% Bayesian Interval with a Triangular Prior Distribution with Pmin = 0.3 or Pmax = 0.7

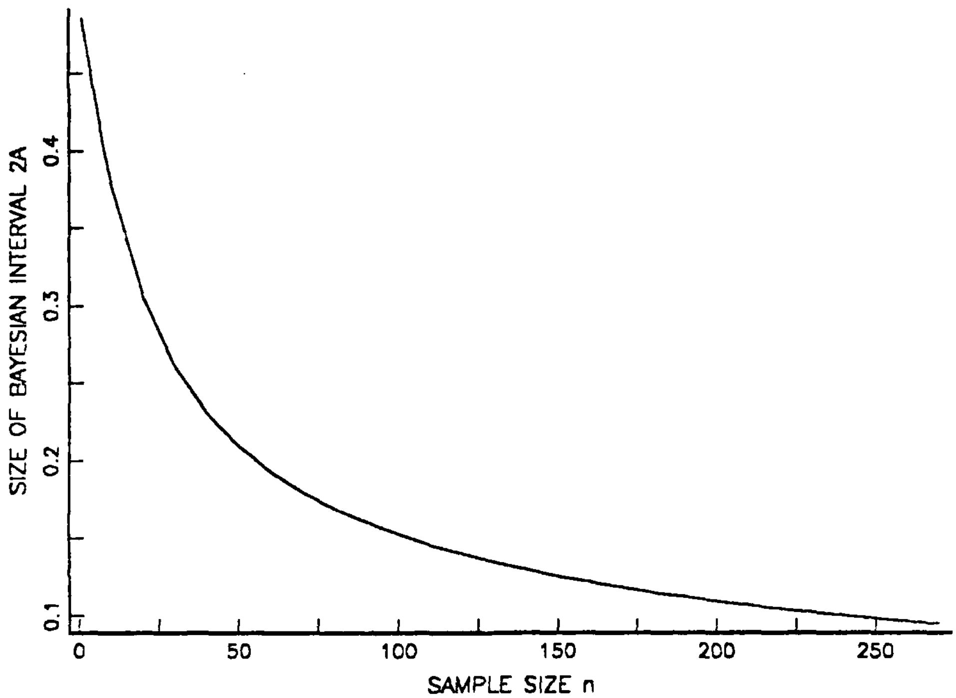

Figure 13. Number of Samples vs The Size of the 95% Bayesian Interval with a Triangular Prior Distribution with Pmin = 0.4 or Pmax = 0.6

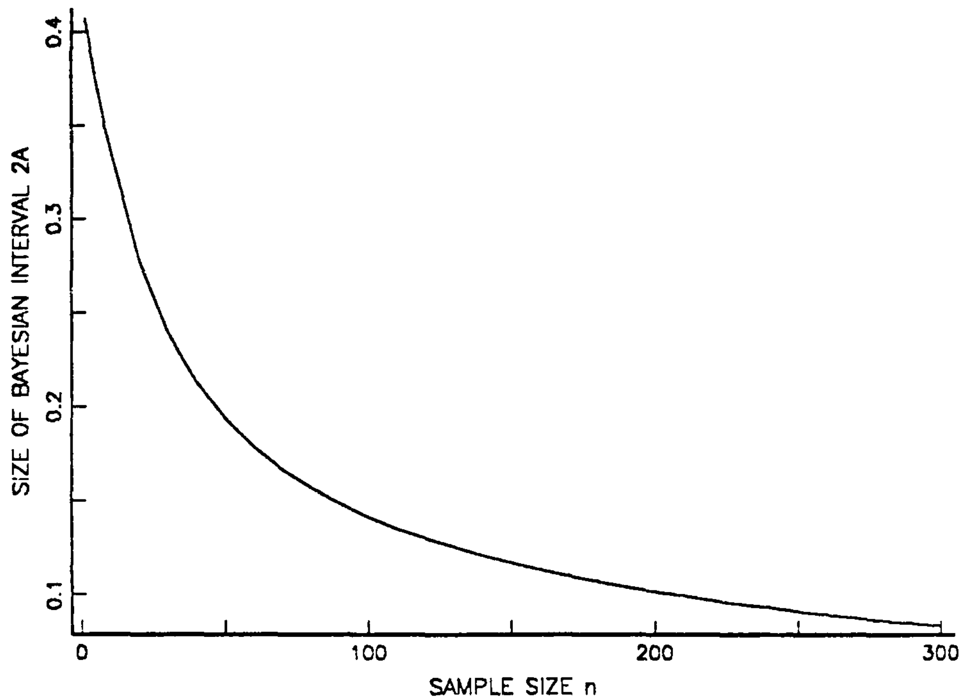

Figure 14. Number of Samples vs The Size of the 95% Bayesian Interval with a Triangular Prior Distribution with Pmin = Pmax = 0.5

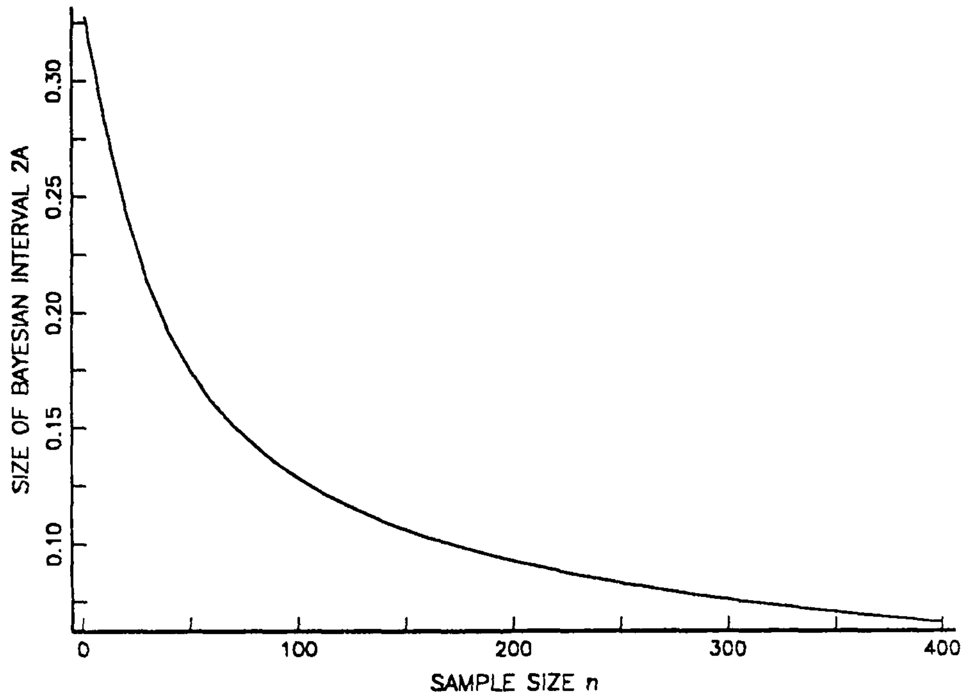

Figure 15. Number of Samples vs The Size of the 95% Bayesian Interval with a Triangular Prior Distribution with Pmin = 0.6 or Pmax = 0.4

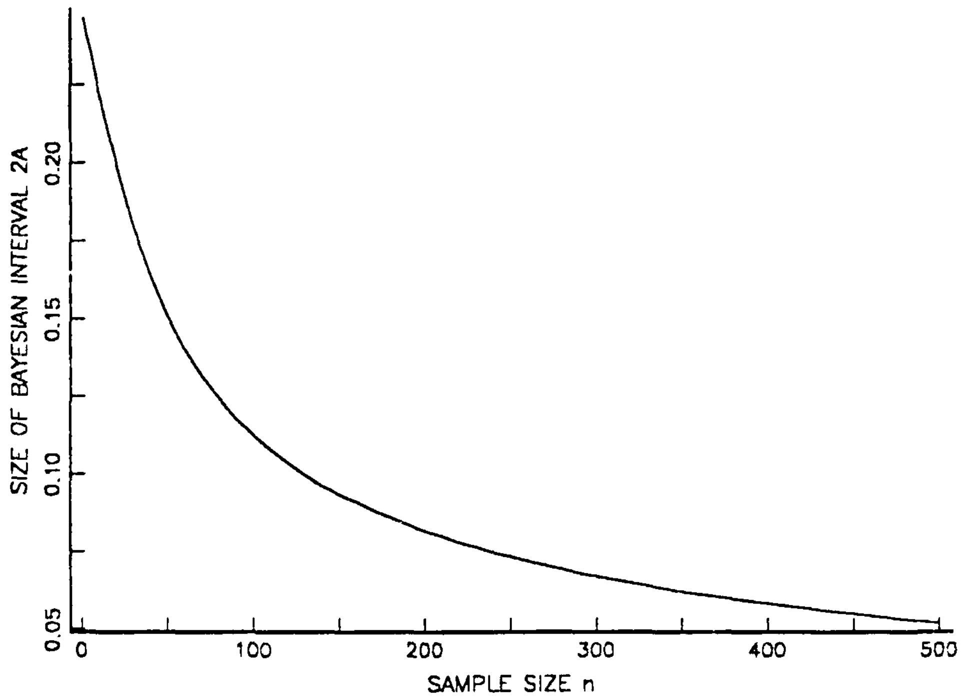

Figure 16. Number of Samples vs The Size of the 95% Bayesian Interval with a Triangular Prior Distribution with Pmin = 0.7 or Pmax = 0.3

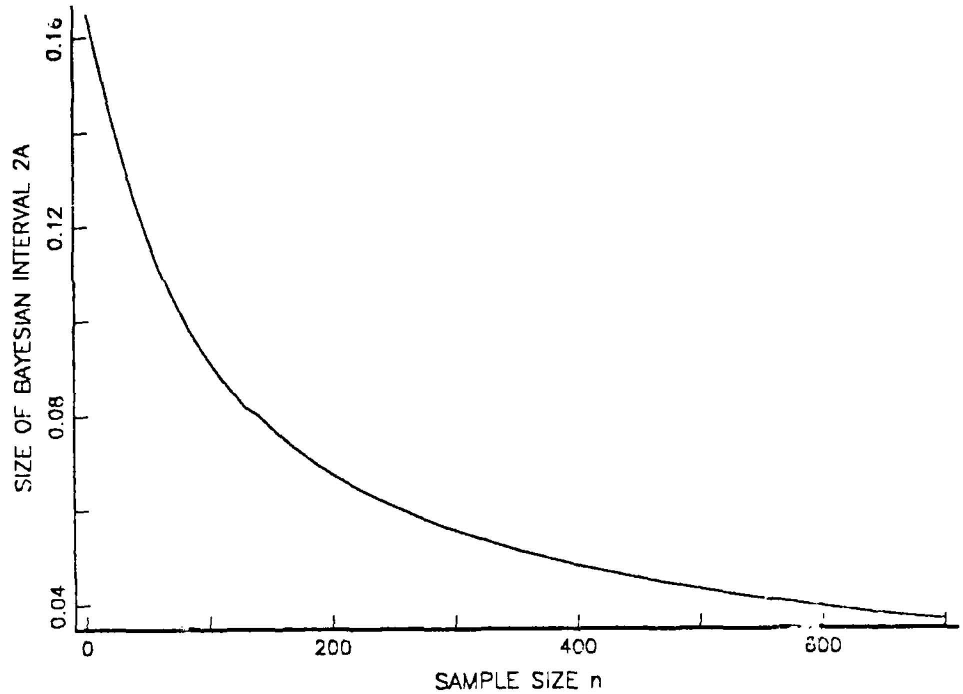

Figure 17. Number of Samples vs The Size of the 95% Bayesian Interval with a Triangular Prior Distribution with Pmin = 0.8 or Pmax = 0.2

LIST OF REFERENCES

1. Larson, Harold J., *Introduction to Probability Theory and Statistical Inference*, John Wiley & Sons, New York. 1982.

2. Manion. Robert B., *Number of Samples Needed to Obtain Desired Bayesian Confidence Intervals for a Proportion.* Master's Thesis. Naval Postgraduate School, Monterey, California, March 1988.

3. Theodore. Floropoulos C.. *A Bayesian Method to Improve Sampling in Weapons Testing.* Master's Thesis, Navai Postgraduate School, Monterey, California. December 1988.

4. Larson. Harold J., *Statistics: An Introduction*. Robert E. Krieger Publishing Company. Inc., Malabar, Florida. 1983.

5. Freund. John E., *Modern Elementary Statistics*. Prentice-Hall. Inc.. Englewood Cliffs. New Jersey. 1984.

6. Berger. James O., *Statistical Decision Theory and Bayesian Analysis*. Springer-Verlag New York Inc., New York. 1980.

7. Jones. Morgan J.. *Introduction to Decision Theory*, Richard D. Irwin. Inc., Homewood. Illinois. 1977.

8. Degroot. Morris H., *Probability and Statistics*, Addison-Wesley Publishing Company. Inc., Massachusetts, 1975.

9. Berkey. Dennis D., *Calculus*. Saunders College Publishing. New York. 1983.

INITIAL DISTRIBUTION LIST

		No. Copies
1.	Defense Technical Information Center Cameron Station Alexandria, VA 22304-6145	2
2.	Library, Code 0142 Naval Postgraduate School Monterey, CA 93943-5002	2
3.	Department of Operations Research Naval Postgraduate School Attn: Prof. Glenn F. Lindsay, Code 55Ls Monterey, CA 93943-5000	2
4.	Department of Operations Research Naval Postgraduate School Attn: Prof. Dan C. Boger, Code 55Bo Monterey, CA 93943-5000	1
5.	Kara Harp Okulu Komutanligi Kutuphanesi Bakanliklar, Ankara TURKEY	1
6.	Bogazici Universitesi Kutuphanesi Istanbul, TURKEY	1
7.	Ortadogu Teknik Universitesi Kutuphanesi Ankara, TURKEY	1
8.	Ahmet Ziyaeddin IPEKKAN Nuri Pamir Caddesi, Koklu Sokak, Guven Apartmani, Daire 7 Kecioren, Ankara, TURKEY	2